

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 16, 2005, 10:10:38 ; Search time 692 Seconds

(without alignments)  
587.535 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 58

Sequence: 1 MGFPPSPPLFFFLQILKQSS.....RYIGKKRRATFRMPRRGTP 58

Scoring table:

OLIGO  
Xgapop 60.0, Xgapext 60.0  
Ygapop 60.0, Ygapext 60.0  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 8765912 seqs, 3504951483 residues

Word size: 1

Total number of hits satisfying chosen parameters: 10478022

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Listing first 45 summaries

Command line parameters: -DEV-xlp  
-O=/cgn2\_1/USPTO\_epool\_p/US10031158/runat\_14102005\_135845\_19714/app\_query.faste\_1.199  
-DB=Published Applications -GFM=fastap -SUFFIX=p2n\_olig\_sz30.rmpb  
-MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTPM=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=30  
-USER=US10031158@cgn2\_1\_1\_722@runat\_14102005\_135845\_19714 -NCPU=6 -ICPU=3  
-NO MAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSELOCK=100 -LONGLOC  
-DEV TIMEOUT=150 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTIS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/2/pubpna/US10J\_NEW\_PUB.seq.\*  
23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	12.1	23	9 US-09-860-996-14	Sequence 14, Appl
2	7	12.1	25	10 US-10-483-654-9	Sequence 9, Appl
3	7	12.1	25	21 US-10-719-900-27595	Sequence 27595, A
4	7	12.1	25	24 US-11-036-317-217607	Sequence 217607, A
5	7	12.1	25	24 US-11-036-317-236358	Sequence 236358, A
6	7	12.1	25	24 US-11-036-317-390331	Sequence 390331, A
7	7	12.1	25	24 US-11-036-317-390331	Sequence 390331, A
8	7	12.1	25	24 US-11-036-317-397073	Sequence 397073, A
9	7	12.1	25	24 US-11-036-317-898261	Sequence 898261, A
10	7	12.1	25	24 US-11-036-317-968427	Sequence 968427, A
11	7	12.1	25	24 US-11-036-317-989913	Sequence 989913, A
12	6	10.3	19	18 US-10-092-900A-680	Sequence 680, App
13	6	10.3	20	17 US-10-181-874-18	Sequence 18, Appl
14	6	10.3	20	17 US-10-271-602B-20	Sequence 20, Appl
15	6	10.3	20	17 US-10-271-602B-29	Sequence 29, Appl
16	6	10.3	20	17 US-10-271-602B-50	Sequence 50, Appl
17	6	10.3	20	17 US-10-349-143-11305	Sequence 11305, A
18	6	10.3	21	20 US-10-473-126-468	Sequence 468, App
19	6	10.3	21	20 US-10-751-736-32850	Sequence 32850, A
20	6	10.3	21	21 US-10-847-918-13631	Sequence 13631, A
21	6	10.3	21	21 US-10-847-918-13633	Sequence 13633, A
22	6	10.3	21	21 US-10-847-918-13634	Sequence 13634, A
23	6	10.3	21	21 US-10-847-918-13635	Sequence 13635, A
24	6	10.3	21	21 US-10-847-918-13639	Sequence 13639, A
25	6	10.3	21	21 US-10-847-918-13640	Sequence 13640, A
26	6	10.3	21	21 US-10-847-918-13641	Sequence 13641, A
27	6	10.3	21	21 US-10-847-918-13642	Sequence 13642, A
28	6	10.3	21	21 US-10-847-918-13644	Sequence 13644, A
29	6	10.3	22	16 US-10-302-379-20	Sequence 0, Appl1
30	6	10.3	23	9 US-09-942-374-6	Sequence 6, Appl1
31	6	10.3	23	18 US-10-665-956-6	Sequence 6, Appl1
32	6	10.3	23	20 US-10-483-654-8	Sequence 8, Appl1
33	6	10.3	25	15 US-10-098-263B-31202	Sequence 31202, A
34	6	10.3	25	15 US-10-098-263B-67759	Sequence 67759, A
35	6	10.3	25	15 US-10-098-263B-78553	Sequence 78553, A
36	6	10.3	25	15 US-10-098-263B-78554	Sequence 78554, A
37	6	10.3	25	15 US-10-098-263B-83546	Sequence 83546, A
38	6	10.3	25	15 US-10-098-263B-121122	Sequence 121122, A
39	6	10.3	25	21 US-10-719-900-45610	Sequence 45610, A
40	6	10.3	25	21 US-10-719-900-46685	Sequence 46685, A
41	6	10.3	25	21 US-10-719-900-46686	Sequence 46686, A
42	6	10.3	25	21 US-10-719-900-47953	Sequence 47953, A
43	6	10.3	25	21 US-10-719-900-124417	Sequence 124417, A
44	6	10.3	25	21 US-10-719-900-174340	Sequence 174340, A
45	6	10.3	25	21 US-10-719-900-241517	Sequence 241517, A

## ALIGNMENTS

RESULT 1  
US-09-860-996-14/c  
; Sequence 14, Application US/09860996  
; Patent No. US2002003433A1  
GENERAL INFORMATION:  
APPLICANT: Mitrophanou, et al  
TITLE OF INVENTION: VECTOR  
FILE REFERENCE: 674523-2010  
CURRENT FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: PCT/GB99/03866  
PRIOR FILING DATE: 1999-11-19  
PRIOR APPLICATION NUMBER: 9825524.3  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn version 3.0

```
; SEQ ID NO 14
; LENGTH: 23
; TYPE: DNA
; ORGANISM: Human Immunodeficiency Virus Type 1
US-09-860-996-14

Alignment Scores:
Pred. No.: 107          Length: 23
Score: 7.00           Matches: 7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 12.07%          Indels: 0
DB: 9                      Gaps: 0

US-10-031-158-14 (1-58) x US-09-860-996-14 (1-23)

QY 5 ProProSerProLeuPhePhe 11
Db 22 CCCCCATCCCCCTTTCTTT 2

RESULT 2
US-10-483-654-9
; Sequence 9, Application US/10483654
; Publication No. US20040220095A1
; GENERAL INFORMATION:
; APPLICANT: Braun, Klaus
; APPLICANT: Waldeck, Waldemar
; APPLICANT: Pipkorn, Rudiger
; APPLICANT: Debus, Jurgen
; APPLICANT: Braun, Isabell
; TITLE OF INVENTION: PNA Conjugate for the Treatment of Diseases Associated with HIV
; FILE REFERENCE: 4121-159
; CURRENT APPLICATION NUMBER: US/10/483,654
; PRIOR FILING DATE: 2004-01-12
; PRIOR APPLICATION NUMBER: PCT/DE 02/02564
; PRIOR FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: DE 101 33 307.2
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 9
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PNA Ia
US-10-483-654-9

Alignment Scores:
Pred. No.: 115          Length: 25
Score: 7.00           Matches: 7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 12.07%          Indels: 0
DB: 20                      Gaps: 0

US-10-031-158-14 (1-58) x US-10-483-654-9 (1-25)

QY 7 SerProLeuPhePhePheLeu 13
Db 4 TCCCCCTTTCTTTTCTTT 24

RESULT 3
US-10-719-900-27595
; Sequence 27595, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; PRIOR FILING DATE: 2003-11-20
; NUMBER OF SEQ ID NOS: 60/427,808
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
```

```
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 27595
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-27595

Alignment Scores:
Pred. No.: 115          Length: 25
Score: 7.00           Matches: 7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 12.07%          Indels: 0
DB: 21                      Gaps: 0

US-10-031-158-14 (1-58) x US-10-719-900-27595 (1-25)

QY 35 LeuMetLeuLeuArgTyrIle 41
Db 5 CTAATGTACTCAGGTATATA 25

RESULT 4
US-11-036-317-217607/c
; Sequence 217607, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Alan
; APPLICANT: Blume, John
; TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
; FILE REFERENCE: 3654.1
; CURRENT APPLICATION NUMBER: US/11/036,317
; PRIOR FILING DATE: 2005-01-13
; PRIOR APPLICATION NUMBER: US 60/536,639
; PRIOR FILING DATE: 2004-01-13
; NUMBER OF SEQ ID NOS: 991174
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 217607
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-036-317-217607

Alignment Scores:
Pred. No.: 115          Length: 25
Score: 7.00           Matches: 7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 12.07%          Indels: 0
DB: 24                      Gaps: 0

US-10-031-158-14 (1-58) x US-11-036-317-217607 (1-25)

QY 32 AsnPheSerLeuMetLeuLeu 38
Db 21 AATTTTCTTGATGTTGTG 1

RESULT 5
US-11-036-317-236358/c
; Sequence 236358, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Alan
; APPLICANT: Blume, John
; TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
; FILE REFERENCE: 3654.1
; CURRENT APPLICATION NUMBER: US/11/036,317
; PRIOR FILING DATE: 2005-01-13
; PRIOR APPLICATION NUMBER: US 60/536,639
; PRIOR FILING DATE: 2004-01-13
; NUMBER OF SEQ ID NOS: 991174
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
```

```
/ SEQ ID NO 236358
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-236358

Alignment Scores:
Pred. No.: 115      Length: 25
Score: 7.00        Matches: 7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 12.07%  Indels: 0
DB: 24             Gaps: 0

US-10-031-158-14 (1-58) x US-11-036-317-236358 (1-25)

Qy 32 AenPhSeRleuMetLeuLen 38
Db 22 AATTTTCTTGATGTGTG 2

RESULT 6
US-11-036-317-358635/c
/ Sequence 358635, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 358635
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-358635

Alignment Scores:
Pred. No.: 115      Length: 25
Score: 7.00        Matches: 7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 12.07%  Indels: 0
DB: 24             Gaps: 0

US-10-031-158-14 (1-58) x US-11-036-317-358635 (1-25)

Qy 32 AenPhSeRleuMetLeuLen 38
Db 23 AATTTTCTTGATGTGTG 3

RESULT 7
US-11-036-317-390331/c
/ Sequence 390331, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 390331
/ LENGTH: 25
/ TYPE: DNA
```

```
/ ORGANISM: Mus musculus
US-11-036-317-390331

Alignment Scores:
Pred. No.: 115      Length: 25
Score: 7.00        Matches: 7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 12.07%  Indels: 0
DB: 24             Gaps: 0

US-10-031-158-14 (1-58) x US-11-036-317-390331 (1-25)

Qy 32 AenPhSeRleuMetLeuLen 38
Db 24 AATTTTCTTGATGTGTG 4

RESULT 8
US-11-036-317-397073/c
/ Sequence 397073, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 397073
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-397073

Alignment Scores:
Pred. No.: 115      Length: 25
Score: 7.00        Matches: 7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 12.07%  Indels: 0
DB: 24             Gaps: 0

US-10-031-158-14 (1-58) x US-11-036-317-397073 (1-25)

Qy 32 AenPhSeRleuMetLeuLen 38
Db 25 AATTTTCTTGATGTGTG 5

RESULT 9
US-11-036-317-898261/c
/ Sequence 898261, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 898261
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-898261
```

Alignment Scores:

Pred. No.:	115	Length:	25
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.07%	Indels:	0
DB:	24	Gaps:	0

US-10-031-158-14 (1-58) x US-11-036-317-989261 (1-25)

QY 32 AaPhSeSerLeuMeLeuLeu 38  
|||||  
DB 22 AATTTTCTTGATGTGTG 2

RESULT 10  
US-11-036-317-968427/c  
Sequence 968427, Application US/11036317  
GENERAL INFORMATION:  
APPLICANT: Williams, Alan  
TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse  
FILE REFERENCE: 3654.1  
CURRENT APPLICATION NUMBER: US/11/036,317  
PRIOR FILING DATE: 2005-01-13  
PRIOR APPLICATION NUMBER: US 60/536,639  
NUMBER OF SEQ ID NOS: 991174  
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1  
SEQ ID NO 968427  
LENGTH: 25  
TYPE: DNA  
ORGANISM: Mus musculus  
US-11-036-317-968427

Alignment Scores:

Pred. No.:	115	Length:	25
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.07%	Indels:	0
DB:	24	Gaps:	0

US-10-031-158-14 (1-58) x US-11-036-317-968427 (1-25)

QY 32 AaPhSeSerLeuMeLeuLeu 38  
|||||  
DB 23 AATTTTCTTGATGTGTG 3

RESULT 11  
US-11-036-317-989913/c  
Sequence 989913, Application US/11036317  
Publication No. US20050214823A1  
GENERAL INFORMATION:  
APPLICANT: Williams, Alan  
APPLICANT: Blume, John  
TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse  
FILE REFERENCE: 3654.1  
CURRENT APPLICATION NUMBER: US/11/036,317  
PRIOR FILING DATE: 2005-01-13  
PRIOR APPLICATION NUMBER: US 60/536,639  
NUMBER OF SEQ ID NOS: 991174  
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1  
SEQ ID NO 989913  
LENGTH: 25  
TYPE: DNA  
ORGANISM: Mus musculus  
US-11-036-317-989913

Alignment Scores:

Pred. No.:	115	Length:	25
Score:	7.00	Matches:	7

Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.07% Indels: 0  
DB: 24 Gaps: 0

US-10-031-158-14 (1-58) x US-11-036-317-989913 (1-25)

QY 32 AaPhSeSerLeuMeLeuLeu 38  
|||||  
DB 25 AATTTTCTTGATGTGTG 5

RESULT 12  
US-10-092-900A-680/c  
Sequence 680, Application US/10092900A  
Publication No. US20040043382A1  
GENERAL INFORMATION:  
APPLICANT: Padigar, Muralidhara  
APPLICANT: Spytek, Kimberly A.  
APPLICANT: Shenoy, Suresh G.  
APPLICANT: Taupier Jr., Raymond J.  
APPLICANT: Pena, Carol E.A.  
APPLICANT: Li, Li  
APPLICANT: Zernusen, Bryan D.  
APPLICANT: Gusev, Vladimir Y.  
APPLICANT: Ji, Weizhen  
APPLICANT: Gorman, Linda  
APPLICANT: Miller, Charles E.  
APPLICANT: Kekuda, Ramesh  
APPLICANT: Paturajan, Meera  
APPLICANT: Gangoli, Bsha A.M.  
APPLICANT: Verne, Corine A.M.  
APPLICANT: Guo, Xiaojia Sasha  
APPLICANT: Tchernev, Velizar T.  
APPLICANT: Fernandes, Elma R.  
APPLICANT: Casman, Stacie J.  
APPLICANT: Malysankar, Uriel M.  
APPLICANT: Gerlach, Valerie  
APPLICANT: Liu, Yi  
APPLICANT: Anderson, David W.  
APPLICANT: Spaderna, Steven K.  
APPLICANT: Catterton, Elina  
APPLICANT: Leite, Mario W.  
APPLICANT: Zhong, Haihong  
APPLICANT: Albrook, John P.  
APPLICANT: Lepley, Denise M.  
APPLICANT: Rieger, Daniel K.  
APPLICANT: Burgess, Catherine E.  
TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 21402-290C  
CURRENT APPLICATION NUMBER: US/10/092,900A  
PRIOR FILING DATE: 2002-03-07  
PRIOR APPLICATION NUMBER: US60/274,322  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: US60/283,675  
PRIOR FILING DATE: 2001-04-13  
PRIOR APPLICATION NUMBER: US60/338,092  
PRIOR FILING DATE: 2001-12-03  
PRIOR APPLICATION NUMBER: US60/274,281  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: US60/274,191  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: US60/325,681  
PRIOR FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: US60/304,354  
PRIOR FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: US60/279,995  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: US60/294,899  
PRIOR FILING DATE: 2001-05-31  
PRIOR APPLICATION NUMBER: US60/287,424  
PRIOR FILING DATE: 2001-04-30  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 768

/ SEQ ID NO 680  
/ LENGTH: 19  
/ TYPE: DNA  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: Description of Artificial Sequence: Forward Primer  
US-10-092-900A-680

Alignment Scores:  
Pred. No.: 934  
Score: 6.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 10.34%  
DB: 18  
Length: 19  
Matches: 6  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-10-031-158-14 (1-58) x US-10-092-900A-680 (1-19)

Qy 48 ThrArgPheTyrAspPro 53  
Db 19 ACTAGATCTCTGGGACCCG 2

RESULT 13  
US-10-181-874-18  
/ Sequence 18, Application US/10181874  
/ Publication No. US20030212020A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Isis Pharmaceuticals, Inc.  
/ APPLICANT: Susan Murray  
/ APPLICANT: Lex M. Cowseart  
/ APPLICANT: Jacqueline Wyatt  
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MACROPHAGE MIGRATION INHIBITORY FACTOR  
/ FILE REFERENCE: RSP-0351  
/ CURRENT APPLICATION NUMBER: US/10181,874  
/ CURRENT FILING DATE: 2002-07-22  
/ PRIOR APPLICATION NUMBER: 09/489,869  
/ PRIOR FILING DATE: 2000-01-20  
/ NUMBER OF SEQ ID NOS: 88  
/ SEQ ID NO 18  
/ LENGTH: 20  
/ TYPE: DNA  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: Antisense Oligonucleotide  
US-10-181-874-18

Alignment Scores:  
Pred. No.: 978  
Score: 6.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 10.34%  
DB: 17  
Length: 20  
Matches: 6  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-10-031-158-14 (1-58) x US-10-181-874-18 (1-20)

Qy 52 AspProArgGlyThr 57  
Db 3 GACCCGCGAGGACACA 20

RESULT 14  
US-10-271-602B-20  
/ Sequence 20, Application US/10271602B  
/ Publication No. US20040002073A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Alice Xiang Li  
/ APPLICANT: Ghazala Hashmi  
/ APPLICANT: Michael Seoul  
/ TITLE OF INVENTION: MULTIPLEXED ANALYSIS OF POLYMORPHIC LOCI  
/ FILE REFERENCE: EMP-US  
/ CURRENT APPLICATION NUMBER: US/10/271,602B  
/ CURRENT FILING DATE: 2002-10-15

/ PRIOR APPLICATION NUMBER: 60/329,427  
/ PRIOR FILING DATE: 2001-10-14  
/ PRIOR APPLICATION NUMBER: 60/329,620  
/ PRIOR FILING DATE: 2001-10-15  
/ PRIOR APPLICATION NUMBER: 60/329,428  
/ PRIOR FILING DATE: 2001-10-14  
/ PRIOR APPLICATION NUMBER: 60/329,619  
/ PRIOR FILING DATE: 2001-10-15  
/ PRIOR APPLICATION NUMBER: 60/364,416  
/ PRIOR FILING DATE: 2002-03-14  
/ NUMBER OF SEQ ID NOS: 212  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 20  
/ LENGTH: 20  
/ TYPE: DNA  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: Probe sequence derived from human genomic sequence  
US-10-271-602B-20

Alignment Scores:  
Pred. No.: 978  
Score: 6.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 10.34%  
DB: 17  
Length: 20  
Matches: 6  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-10-031-158-14 (1-58) x US-10-271-602B-20 (1-20)

Qy 12 PheLeuGlnLeuLeuYs 17  
Db 1 TTCTTGACGCTCTTAAG 18

RESULT 15  
US-10-271-602B-29  
/ Sequence 29, Application US/10271602B  
/ Publication No. US20040002073A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Alice Xiang Li  
/ APPLICANT: Ghazala Hashmi  
/ APPLICANT: Michael Seoul  
/ TITLE OF INVENTION: MULTIPLEXED ANALYSIS OF POLYMORPHIC LOCI  
/ FILE REFERENCE: EMP-US  
/ CURRENT APPLICATION NUMBER: US/10/271,602B  
/ CURRENT FILING DATE: 2002-10-15  
/ PRIOR APPLICATION NUMBER: 60/329,427  
/ PRIOR FILING DATE: 2001-10-14  
/ PRIOR APPLICATION NUMBER: 60/329,620  
/ PRIOR FILING DATE: 2001-10-15  
/ PRIOR APPLICATION NUMBER: 60/329,428  
/ PRIOR FILING DATE: 2001-10-14  
/ PRIOR APPLICATION NUMBER: 60/329,619  
/ PRIOR FILING DATE: 2001-10-15  
/ PRIOR APPLICATION NUMBER: 60/364,416  
/ PRIOR FILING DATE: 2002-03-14  
/ NUMBER OF SEQ ID NOS: 212  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 29  
/ LENGTH: 20  
/ TYPE: DNA  
/ ORGANISM: Human  
US-10-271-602B-29

Alignment Scores:  
Pred. No.: 978  
Score: 6.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 10.34%  
DB: 17  
Length: 20  
Matches: 6  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 16, 2005, 09:22:02 ; Search time 130 Seconds  
(without alignments)  
730.031 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 58

Sequence: 1 MOMPSPSPFLPFLQLKQSS.....RYGKKRRATRFWDPRGRTP 58

Scoring table: OLIGO  
Xgapop 60.0, Xgapext 60.0  
Ygapop 60.0, Ygapext 60.0  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Word size: 1

Total number of hits satisfying chosen parameters: 965324

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: listing first 45 summaries

Command line parameters:

-MODX=frame+ p2n model -DEV=rlp  
-Q/cg2\_1/USPTO.epool\_p/US10031158/rnat\_14102005\_135843\_19628/app\_query.faeta\_1.199  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=p2n\_olig\_bz30.rml -MINMATCH=0.1  
-LOOPEXT=0 -UNITS=bases -START=1 -END=1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=30  
-USER=US10031158 @CCN 1.1.105 @rnat\_14102005\_135843\_19628 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEBUFFER -NEG\_SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

- 1: Issued Patents NA.\*
- 2: /cg2\_6/prodata/1/ina/5A\_COMB.seq.\*
- 3: /cg2\_6/prodata/1/ina/5B\_COMB.seq.\*
- 4: /cg2\_6/prodata/1/ina/6A\_COMB.seq.\*
- 5: /cg2\_6/prodata/1/ina/6B\_COMB.seq.\*
- 6: /cg2\_6/prodata/1/ina/PCTUS\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	10.3	20	3	US-09-489-869-18
2	6	10.3	20	4	US-09-422-978-11305
3	6	10.3	22	4	US-08-857-636-20
4	6	10.3	25	4	US-09-396-196G-41485
5	6	10.3	25	4	US-09-396-196G-41486
6	6	10.3	25	4	US-09-396-196G-41487
7	6	10.3	25	4	US-09-396-196G-41488
8	6	10.3	25	4	US-09-396-196G-41489
9	6	10.3	25	4	US-09-396-196G-41490
10	6	10.3	25	4	US-09-396-196G-113910
11	6	10.3	25	4	US-09-396-196G-127372
12	6	10.3	25	4	US-09-396-196G-127373

13	6	10.3	25	4	US-09-396-196G-127384	Sequence 127384,
C 14	6	10.3	30	1	US-08-186-229-20	Sequence 20, Appl
C 15	6	10.3	30	2	US-08-470-124-20	Sequence 20, Appl
C 16	5	8.6	16	1	US-08-087-387-1	Sequence 1, Appl1
C 17	5	8.6	16	1	US-08-455-627-1	Sequence 1, Appl1
C 18	5	8.6	16	1	US-08-242-664-21	Sequence 21, Appl
C 19	5	8.6	16	1	US-08-242-664-22	Sequence 22, Appl
C 20	5	8.6	16	1	US-08-484-138-21	Sequence 21, Appl
C 21	5	8.6	16	1	US-08-484-138-22	Sequence 22, Appl
C 22	5	8.6	16	1	US-08-461-271-1	Sequence 1, Appl1
C 23	5	8.6	16	1	US-08-713-685A-1	Sequence 1, Appl1
C 24	5	8.6	16	2	US-08-689-856-1	Sequence 1, Appl1
C 25	5	8.6	16	2	US-08-412-376-1	Sequence 1, Appl1
C 26	5	8.6	16	3	US-09-070-477-1	Sequence 1, Appl1
C 27	5	8.6	16	3	US-07-808-452-17	Sequence 17, Appl
C 28	5	8.6	16	3	US-09-133-717-11	Sequence 11, Appl
C 29	5	8.6	16	3	US-08-787-321-1	Sequence 1, Appl1
C 30	5	8.6	16	3	US-08-787-321-16	Sequence 16, Appl
C 31	5	8.6	16	3	US-09-158-863C-11	Sequence 11, Appl
C 32	5	8.6	16	5	PCT-US92-10770-17	Sequence 17, Appl
C 33	5	8.6	16	5	PCT-US92-10792-15	Sequence 15, Appl
C 34	5	8.6	16	5	PCT-US95-06379-21	Sequence 21, Appl
C 35	5	8.6	16	5	PCT-US95-06379-22	Sequence 22, Appl
C 36	5	8.6	17	1	US-08-105-483-85	Sequence 85, Appl
C 37	5	8.6	17	1	US-08-709-209-85	Sequence 85, Appl
C 38	5	8.6	17	1	US-08-458-101-85	Sequence 85, Appl
C 39	5	8.6	17	3	US-09-049-277B-17	Sequence 17, Appl
C 40	5	8.6	17	3	US-08-584-040-2777	Sequence 2777, Ap
C 41	5	8.6	17	3	US-08-584-040-2778	Sequence 2778, Ap
C 42	5	8.6	17	3	US-08-584-040-2779	Sequence 2779, Ap
C 43	5	8.6	17	3	US-08-679-645-850	Sequence 850, App
C 44	5	8.6	17	3	US-08-679-645-852	Sequence 852, App
C 45	5	8.6	17	4	US-09-371-772B-1301	Sequence 1301, Ap

#### ALIGNMENTS

RESULT 1  
US-09-489-869-18  
Sequence 18, Application US/09489869A  
Patent No. 6268151  
GENERAL INFORMATION:  
APPLICANT: Susan Murray  
APPLICANT: Lex M. Cowart  
APPLICANT: Jacqueline Wyatt  
TITLE OF INVENTION: ANTISENSE MODULATION OF MACROPHAGE MIGRATION INHIBITORY FACTOR  
FILE REFERENCE: RTS-0110  
CURRENT APPLICATION NUMBER: US/09/489,869A  
CURRENT FILING DATE: 2000-01-20  
NUMBER OF SEQ ID NOS: 88  
SEQ ID NO 18  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Antisense Oligonucleotide  
US-09-489-869-18

#### Alignment Scores:

Pred. No.: 178  
Score: 6.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 10.34%  
DB: 3  
Gaps: 0

US-10-031-158-14 (1-58) x US-09-489-869-18 (1-20)

OY 52 AAPPPOARGARGG1YThr 57  
DB 3 GACCCGCGACAGGACACA 20

RESULT 2

US-09-422-978-11305  
Sequence 11305, Application US/09422978  
Patent No. 6537751  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Chumakov, Ilya  
TITLE OF INVENTION: Biallelic markers for use in constructing a high density...  
FILE REFERENCE: GENSET.020CP1  
CURRENT APPLICATION NUMBER: US/09/422,978  
CURRENT FILING DATE: 1999-10-20  
EARLIER APPLICATION NUMBER: US 09/298,850  
EARLIER FILING DATE: 1999-04-21  
EARLIER APPLICATION NUMBER: US 60/109,732  
EARLIER FILING DATE: 1998-11-23  
EARLIER APPLICATION NUMBER: US 60/082,614  
EARLIER FILING DATE: 1998-04-21  
NUMBER OF SEQ ID NOS: 11796  
SEQ ID NO 11305  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: primer\_bind  
LOCATION: 1..20  
OTHER INFORMATION: downstream amplification primer 99-4119 for SEQ 3440, in compleme  
US-09-422-978-11305

Alignment Scores:  
Pred. No.: 178 Length: 20  
Score: 6.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.34% Indels: 0  
Gaps: 0  
DB: 4

US-10-031-158-14 (1-58) X US-09-422-978-11305 (1-20)

Qy 8 ProLeuphaphelenu 13  
|||||  
DB 3 CCCCTATTTTCTCTG 20

RESULT 3  
US-08-857-636-20/c

Sequence 20, Application US/08857636  
Patent No. 6552181

GENERAL INFORMATION:

APPLICANT: Dean, Michael Carlton  
APPLICANT: Hann, Heidi Eve  
APPLICANT: Wickling, Carol  
APPLICANT: Christiansen, Jeffrey  
APPLICANT: Zaphiropoulos, Peter G.  
APPLICANT: Gallant, Mae R.  
APPLICANT: Shanley, Susan Mary  
APPLICANT: Chidambaram, Abirami  
APPLICANT: Vorechovsky, Igor  
APPLICANT: Holmberg-Lindstrom, Erika  
APPLICANT: Unden, Anne Birgitte  
APPLICANT: Gillies, Susan Alana  
APPLICANT: Negus, Kylie  
APPLICANT: Smyth, Ian Mcleod  
APPLICANT: Pressman, Carol Leah  
APPLICANT: Leftehl, David J.  
APPLICANT: Gerrard, Bernard  
APPLICANT: Goldstein, Alisa Miriam  
APPLICANT: Mainwright, Brandon  
APPLICANT: Toftgard, Rune Carl-Magnus  
APPLICANT: Chenevix-Trench, Georgia  
APPLICANT: Bale, Allen E.  
TITLE OF INVENTION: A Basal Cell Carcinoma Tumor Suppressor Gene  
FILE REFERENCE: 3101.1  
NUMBER OF SEQUENCES: 83  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/857,636  
FILING DATE: 16-MAY-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/017,906  
FILING DATE: 17-MAY-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU P00011  
FILING DATE: 21-MAY-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU P00363  
FILING DATE: 07-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/019,765  
FILING DATE: 14-JUN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunter, Tom  
REGISTRATION NUMBER: 38,498  
REFERENCE/DOCKET NUMBER: 015280-278200US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: -  
LOCATION: 1..22  
OTHER INFORMATION: /note= "PTCF21 primer"  
US-08-857-636-20

Alignment Scores:

Pred. No.: 195 Length: 22  
Score: 6.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.34% Indels: 0  
Gaps: 0  
DB: 4

US-10-031-158-14 (1-58) X US-08-857-636-20 (1-22)

Qy 28 ValpheleuArgaAenPhe 33  
|||||  
DB 21 GGTTCCTGAGAAATTT 4

RESULT 4  
US-09-396-196G-41485/c

Sequence 41485, Application US/09396196G  
Patent No. 6821724

GENERAL INFORMATION:

APPLICANT: Michael Miltmann  
APPLICANT: David Mack  
APPLICANT: David Lockhart  
APPLICANT: Affymetrix, Inc.  
TITLE OF INVENTION: Methods of Genetic Analysis  
FILE REFERENCE: 3101.1  
CURRENT APPLICATION NUMBER: US/09/396,196G  
CURRENT FILING DATE: 1999-09-15



```
/ PRIOR APPLICATION NUMBER: 60/100,678
/ PRIOR FILING DATE: 1998-09-17
/ NUMBER OF SEQ ID NOS: 127806
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 41485
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-09-396-196G-41485

Alignment Scores:
Pred. No.: 220
Score: 6.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 10.34%
DB: 4
Gaps: 0

US-10-031-158-14 (1-58) x US-09-396-196G-41485 (1-25)

QY 46 ArgAlaThrArgPheTIP 51
DB 24 CGTGTACACAGGTTCTGG 7

RESULT 5
US-09-396-196G-41486/C
/ Sequence 41486, Application US/09396196G
/ Patent No. 6821724
/ GENERAL INFORMATION:
/ APPLICANT: Michael Miltmann
/ APPLICANT: David Mack
/ APPLICANT: David Lockhart
/ APPLICANT: Affymetrix, Inc.
/ TITLE OF INVENTION: Methods of Genetic Analysis
/ FILE REFERENCE: 3101.1
/ CURRENT APPLICATION NUMBER: US/09/396,196G
/ PRIOR FILING DATE: 1999-09-15
/ PRIOR APPLICATION NUMBER: 60/100,678
/ PRIOR FILING DATE: 1998-09-17
/ NUMBER OF SEQ ID NOS: 127806
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 41486
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-09-396-196G-41486

Alignment Scores:
Pred. No.: 220
Score: 6.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 10.34%
DB: 4
Gaps: 0

US-10-031-158-14 (1-58) x US-09-396-196G-41486 (1-25)

QY 46 ArgAlaThrArgPheTIP 51
DB 23 CGTGTACACAGGTTCTGG 6

RESULT 6
US-09-396-196G-41487/C
/ Sequence 41487, Application US/09396196G
/ Patent No. 6821724
/ GENERAL INFORMATION:
/ APPLICANT: Michael Miltmann
/ APPLICANT: David Mack
/ APPLICANT: David Lockhart
/ APPLICANT: Affymetrix, Inc.
/ TITLE OF INVENTION: Methods of Genetic Analysis
/ FILE REFERENCE: 3101.1
/ CURRENT APPLICATION NUMBER: US/09/396,196G
```

```
/ CURRENT FILING DATE: 1999-09-15
/ PRIOR APPLICATION NUMBER: 60/100,678
/ PRIOR FILING DATE: 1998-09-17
/ NUMBER OF SEQ ID NOS: 127806
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 41487
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-09-396-196G-41487

Alignment Scores:
Pred. No.: 220
Score: 6.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 10.34%
DB: 4
Gaps: 0

US-10-031-158-14 (1-58) x US-09-396-196G-41487 (1-25)

QY 46 ArgAlaThrArgPheTIP 51
DB 22 CGTGTACACAGGTTCTGG 5

RESULT 7
US-09-396-196G-41488/C
/ Sequence 41488, Application US/09396196G
/ Patent No. 6821724
/ GENERAL INFORMATION:
/ APPLICANT: Michael Miltmann
/ APPLICANT: David Mack
/ APPLICANT: David Lockhart
/ APPLICANT: Affymetrix, Inc.
/ TITLE OF INVENTION: Methods of Genetic Analysis
/ FILE REFERENCE: 3101.1
/ CURRENT APPLICATION NUMBER: US/09/396,196G
/ PRIOR FILING DATE: 1999-09-15
/ PRIOR APPLICATION NUMBER: 60/100,678
/ PRIOR FILING DATE: 1998-09-17
/ NUMBER OF SEQ ID NOS: 127806
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 41488
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-09-396-196G-41488

Alignment Scores:
Pred. No.: 220
Score: 6.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 10.34%
DB: 4
Gaps: 0

US-10-031-158-14 (1-58) x US-09-396-196G-41488 (1-25)

QY 46 ArgAlaThrArgPheTIP 51
DB 21 CGTGTACACAGGTTCTGG 4

RESULT 8
US-09-396-196G-41489/C
/ Sequence 41489, Application US/09396196G
/ Patent No. 6821724
/ GENERAL INFORMATION:
/ APPLICANT: Michael Miltmann
/ APPLICANT: David Mack
/ APPLICANT: David Lockhart
/ APPLICANT: Affymetrix, Inc.
/ TITLE OF INVENTION: Methods of Genetic Analysis
/ FILE REFERENCE: 3101.1
```

```
/ CURRENT APPLICATION NUMBER: US/09/396,196G
/ CURRENT FILING DATE: 1999-09-15
/ PRIOR APPLICATION NUMBER: 60/100,678
/ PRIOR FILING DATE: 1998-09-17
/ NUMBER OF SEQ ID NOS: 127806
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 41489
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-09-396-196G-41489

Alignment Scores:
Pred. No.: 220
Score: 6.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 10.34%
DB: 4
Gaps: 0

US-10-031-158-14 (1-58) x US-09-396-196G-41489 (1-25)

QY 46 ArgAlaThrArgPheTyr 51
DB 19 CGTCTACAGGTTCTGG 2

RESULT 9
US-09-396-196G-41490/c
/ Sequence 41490, Application US/09396196G
/ Patent No. 6821724
/ GENERAL INFORMATION:
/ APPLICANT: Michael Miltmann
/ APPLICANT: David Mack
/ APPLICANT: David Lockhart
/ APPLICANT: Affymetrix, Inc.
/ TITLE OF INVENTION: Methods of Genetic Analysis
/ FILE REFERENCE: 3101.1
/ CURRENT APPLICATION NUMBER: US/09/396,196G
/ CURRENT FILING DATE: 1999-09-15
/ PRIOR APPLICATION NUMBER: 60/100,678
/ PRIOR FILING DATE: 1998-09-17
/ NUMBER OF SEQ ID NOS: 127806
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 41490
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-09-396-196G-41490

Alignment Scores:
Pred. No.: 220
Score: 6.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 10.34%
DB: 4
Gaps: 0

US-10-031-158-14 (1-58) x US-09-396-196G-41490 (1-25)

QY 46 ArgAlaThrArgPheTyr 51
DB 18 CGTCTACAGGTTCTGG 1

RESULT 10
US-09-396-196G-113910/c
/ Sequence 113910, Application US/09396196G
/ Patent No. 6821724
/ GENERAL INFORMATION:
/ APPLICANT: Michael Miltmann
/ APPLICANT: David Mack
/ APPLICANT: David Lockhart
/ APPLICANT: Affymetrix, Inc.
/ TITLE OF INVENTION: Methods of Genetic Analysis
```

```
/ FILE REFERENCE: 3101.1
/ CURRENT APPLICATION NUMBER: US/09/396,196G
/ CURRENT FILING DATE: 1999-09-15
/ PRIOR APPLICATION NUMBER: 60/100,678
/ PRIOR FILING DATE: 1998-09-17
/ NUMBER OF SEQ ID NOS: 127806
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 113910
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-09-396-196G-113910

Alignment Scores:
Pred. No.: 220
Score: 6.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 10.34%
DB: 4
Gaps: 0

US-10-031-158-14 (1-58) x US-09-396-196G-113910 (1-25)

QY 18 GlnSerArgArgLeu 23
DB 23 CAAAGCAGCCGTCGCTT 6

RESULT 11
US-09-396-196G-127372
/ Sequence 127372, Application US/09396196G
/ Patent No. 6821724
/ GENERAL INFORMATION:
/ APPLICANT: Michael Miltmann
/ APPLICANT: David Mack
/ APPLICANT: David Lockhart
/ APPLICANT: Affymetrix, Inc.
/ TITLE OF INVENTION: Methods of Genetic Analysis
/ FILE REFERENCE: 3101.1
/ CURRENT APPLICATION NUMBER: US/09/396,196G
/ CURRENT FILING DATE: 1999-09-15
/ PRIOR APPLICATION NUMBER: 60/100,678
/ PRIOR FILING DATE: 1998-09-17
/ NUMBER OF SEQ ID NOS: 127806
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 127372
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-09-396-196G-127372

Alignment Scores:
Pred. No.: 220
Score: 6.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 10.34%
DB: 4
Gaps: 0

US-10-031-158-14 (1-58) x US-09-396-196G-127372 (1-25)

QY 19 SerSerArgArgLeu 24
DB 7 AGCAGTCGACAGCTCGAA 24

RESULT 12
US-09-396-196G-127373
/ Sequence 127373, Application US/09396196G
/ Patent No. 6821724
/ GENERAL INFORMATION:
/ APPLICANT: Michael Miltmann
/ APPLICANT: David Mack
/ APPLICANT: David Lockhart
/ APPLICANT: Affymetrix, Inc.
```

;; TITLE OF INVENTION: Methods of Genetic Analysis  
;; FILE REFERENCE: 3101.1  
;; CURRENT APPLICATION NUMBER: US/09/396,196G  
;; CURRENT FILING DATE: 1999-09-15  
;; PRIOR APPLICATION NUMBER: 60/100,678  
;; PRIOR FILING DATE: 1998-09-17  
;; NUMBER OF SEQ ID NOS: 127806  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO: 127373  
;; LENGTH: 25  
;; TYPE: DNA  
;; ORGANISM: mus musculus  
US-09-396-196G-127373

Alignment Scores:  
Pred. No.: 220 Length: 25  
Score: 6.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.34% Indels: 0  
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-396-196G-127373 (1-25)

OY 19 SerSerArgArgLeuGlu 24  
DB 4 AGCAGTCGACAGCTCGAA 21

RESULT 13  
US-09-396-196G-127384  
; Sequence 127384, Application US/09396196G  
; Patent No. 6821724  
; GENERAL INFORMATION:  
; APPLICANT: Michael Miltmann  
; APPLICANT: David Mack  
; APPLICANT: David Lockhart  
; APPLICANT: Affymetrix, Inc.  
; TITLE OF INVENTION: Methods of Genetic Analysis  
; FILE REFERENCE: 3101.1  
; CURRENT APPLICATION NUMBER: US/09/396,196G  
; CURRENT FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: 60/100,678  
; PRIOR FILING DATE: 1998-09-17  
; NUMBER OF SEQ ID NOS: 127806  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 127384  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: mus musculus  
US-09-396-196G-127384

Alignment Scores:  
Pred. No.: 220 Length: 25  
Score: 6.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.34% Indels: 0  
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-396-196G-127384 (1-25)

OY 19 SerSerArgArgLeuGlu 24  
DB 1 AGCAGTCGACAGCTCGAA 18

RESULT 14  
US-08-186-229-20/c  
; Sequence 20, Application US/08186229  
; Patent No. 5736316  
; GENERAL INFORMATION:  
; APPLICANT: Irvine, Bruce D.  
; APPLICANT: Kolberg, Janice A.  
; APPLICANT: Running, Joyce A.

;; APPLICANT: Urdea, Michael S.  
;; TITLE OF INVENTION: HBV PROBES FOR USE IN SOLUTION  
;; TITLE OF INVENTION: PHASE SANDWICH HYBRIDIZATION ASSAYS  
;; NUMBER OF SEQUENCES: 55  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Morrison & Foerster  
;; STREET: 755 Page Mill Road  
;; CITY: Palo Alto  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/186,229  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/813,586  
FILING DATE: 23-DEC-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas E. Clotf1  
REGISTRATION NUMBER: 21, 013  
REFERENCE/DOCKET NUMBER: 22300-20234.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-813-5600  
TELEFAX: 415-494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-186-229-20

Alignment Scores:  
Pred. No.: 261 Length: 30  
Score: 6.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.34% Indels: 0  
DB: 1 Gaps: 0

US-10-031-158-14 (1-58) x US-08-186-229-20 (1-30)

OY 18 GlnSerSerArgArgLeu 23  
DB 21 CAGTCTCGAGAGATTG 4

RESULT 15  
US-08-470-124-20/c  
; Sequence 20, Application US/08470124  
; Patent No. 5649481  
; GENERAL INFORMATION:  
; APPLICANT: Urdea, Michael S.  
; APPLICANT: Horn, Thomas  
; APPLICANT: Chang, Chu-An  
; APPLICANT: Warner, Brian  
; APPLICANT: Pultz, Timothy J.  
; TITLE OF INVENTION: LARGE COMB-TYPE BRANCHED  
; TITLE OF INVENTION: POLYNUCLEOTIDES  
; NUMBER OF SEQUENCES: 87  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 545 Middlefield Road, Suite 200  
CITY: Menlo Park  
STATE: California  
COUNTRY: USA  
ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/470,124  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/813,588  
 FILING DATE: 23 December 1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ciotti, Thomas E.  
 REGISTRATION NUMBER: 21,013  
 REFERENCE/DOCKET NUMBER: 22300-20104.20  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-813-5600  
 TELEFAX: 415-327-2951  
 TELEX: 706141  
 INFORMATION FOR SEQ ID NO: 20:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 30 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-470-124-20

Alignment Scores:

Pred. No.:	261	Length:	30
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x US-08-470-124-20 (1-30)

Qy	18	GinSerSerArgArgLeu	23
Db	21	CAGTCCTCGMGRAGATTG	4

Search completed: October 16, 2005, 11:02:49  
 Job time : 131 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 16, 2005, 08:49:13 ; Search time 2981 Seconds

(without alignments)  
740.600 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 58

Sequence: 1 MOMPFPPLFFFLQLKQSS.....RYGKRRATFMDPRGTP 58

Scoring table:

OLIGO	Xgapop 60.0	Xgapext 60.0
	Ygapop 60.0	Ygapext 60.0
	Fgapop 6.0	Fgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 34239544 segs, 19032134700 residues

Word size: 1

Total number of hits satisfying chosen parameters: 45551

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=xlp  
-O=/sgn2.1/USPFO\_epool\_p/US10031158/runat\_14102005\_135843\_19614/app\_query.fasta\_1.199  
-DB=EST -QWNT=fastap -SUFFIX=p2n\_olig\_gz30.rst -MINMATCH=0.1 -LOOKCL=0  
-LOOKEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=30  
-USER=US10031158 @CGN 1.1 5180 @runat\_14102005\_135843\_19614 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEOQUERY -NEG SCORES=0 -WAIT -DSEBLOCK=100 -LONGLOC  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST :  
1: gb\_est1 :  
2: gb\_est2 :  
3: gb\_hic :  
4: gb\_est3 :  
5: gb\_est4 :  
6: gb\_est5 :  
7: gb\_est6 :  
8: gb\_gse1 :  
9: gb\_gse2 :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	10.3	28	AZ87839	1M0147H22
2	5	8.6	16	BQ583458	E011979-0
3	5	8.6	19	AZ465954	1M0276B16
4	5	8.6	20	AW246466	2821777.3
5	5	8.6	20	CF282165	14EUTL-09
6	5	8.6	20	AZ832043	2M0112111
7	5	8.6	20	AZ957966	2M0225101
8	5	8.6	20	AG189044	Pan tlog1
9	5	8.6	20	CL660020	PR10135d_

10	5	8.6	21	CF338234	RCL1--01-
11	5	8.6	21	AZ31619	1M0059X09
12	5	8.6	21	AZ341108	AZ341108
13	5	8.6	21	AZ589400	1M0398C23
14	5	8.6	21	AZ806282	2M0068A15
15	5	8.6	21	AZ850337	AZ850337
16	5	8.6	21	AZ943299	2M0203X21
17	5	8.6	22	A1688330	2M0444C14
18	5	8.6	22	A1661940	2M0089P23
19	5	8.6	22	CF310806	ABF-05-K
20	5	8.6	22	AZ317017	1M0035P09
21	5	8.6	22	TA231E08Q	T. brucei
22	5	8.6	23	CF302431	7LEAF--07
23	5	8.6	23	AZ615086	AZ615086
24	5	8.6	23	AZ819376	2M0089P23
25	5	8.6	23	TA172005P	TA172005P
26	5	8.6	24	AU257964	AU257964
27	5	8.6	24	CF295238	30DGS--05
28	5	8.6	24	CF340367	RCL1--07
29	5	8.6	24	AZ852116	2M0154M12
30	5	8.6	24	TA155F120	TA155F120
31	5	8.6	24	TA185C06P	TA185C06P
32	5	8.6	25	A1015571	OV09F01.X
33	5	8.6	25	A1624473	LS6Bd11.X
34	5	8.6	25	AJ663467	AJ663467
35	5	8.6	25	AZ997733	2M0284E22
36	5	8.6	25	BH854637	KG07068-5
37	5	8.6	25	TA232G09Q	TA232G09Q
38	5	8.6	26	AJ666261	AJ666261
39	5	8.6	26	AJ683031	AJ683031
40	5	8.6	26	AJ685374	AJ685374
41	5	8.6	26	AJ665518	AJ665518
42	5	8.6	26	AZ377143	1M0131017
43	5	8.6	26	AZ447827	1M0245J08
44	5	8.6	26	TA319603Q	TA319603Q
45	5	8.6	27	BQ592207	E012696-0

## ALIGNMENTS

RESULT 1  
AZ87839 28 bp DNA linear GSS 02-OCT-2000  
1M0147H22R Mouse 10kb plasmid UGCM1 library Mus musculus genomic  
LOCUS  
DEFINITION  
1M0147H22R Mouse 10kb plasmid UGCM1 library Mus musculus genomic  
ACCESSION  
AZ87839.1 GI:10501547  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
GSS.  
Mus musculus  
Mus musculus  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 28)  
Dunn, D., Aoyagi, A., Barber, M., Baconn, T., Duval, B., Hamill, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,  
Niederhauser, A. and Wright, D., Weiser, R., Tingey, A., von  
Plasmid inserts  
Mouse whole genome scaffolding with paired end reads from 10kb  
unpublished (2000)  
Contact: Robert B. Weiser  
University of Utah  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert length: 10000 Std Error: 0.00  
Plate: 0147 row: H column: 22  
Seq primer: CACACAGGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 28.

JOURNAL  
COMMENT

## FEATURES

## SOURCE

## Location/Qualifiers

1. 28  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="U081M0147H23"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_1lb="Mouse 10kb plasmid U081M library"  
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/anases/). The DNA was hydrodynamically sheared by repeated passages through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g1|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

## Alignment Scores:

Pred. No.: 5.35e+03 Length: 28  
Score: 6.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.34% Indels: 0  
DB: 8 Gaps: 0

US-10-031-158-14 (1-58) x AZ387839 (1-28)

OY 8 Proleuphaphelenu 13

DB 11 CCATATTTTCTTCTT 28

## RESULT 2

BQ583458 16 bp mRNA linear EST 06-DEC-2002

LOCUS B011979-024-005-011-SP6 MP1Z-ADIS-024-inflorance Beta vulgaris

DEFINITION CDNA clone 024-005-011 5-PRIME, mRNA sequence.

ACCESSION BQ583458

VERSION BQ583458.1 GI:26113035

KEYWORDS EST.

SOURCE Beta vulgaris

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE 1 (bases 1 to 16)

AUTHORS Herwig, R., Schulz, B., Weisshaar, B., Hennig, S., Steinfath, M.,

TITLE and Radloff, U.

JOURNAL Construction of a 'unigene' cDNA clone set by oligonucleotide

MEDLINE fingerprinting allows access to 25 000 potential sugar beet genes

COMMENT Plant J. 32 (5), 845-857 (2002)

ADIS DNA core facility at MP1Z

Max-Planck-Institute for Plant Breeding Research

Carl-von-Linne Weg 10, 50829 Koeln, Germany

## FEATURES

## SOURCE

Insert Length: 16 Std Error: 0.00  
Plate: 5 row: J column: 11  
Seq primer: SP6; CAPACATTTAGTGACTATAG.  
Location/Qualifiers

1. 16  
/organism="Beta vulgaris"  
/mol\_type="mRNA"  
/cultivar="KWS2320 (double haploid, monogerm breeding line)"  
/db\_xref="GABI:183240"  
/db\_xref="taxon:161934"  
/clone="024-005-011"  
/tissue\_type="inflorescence"  
/lab\_host="EMDH108"  
/clone\_1lb="MP1Z-ADIS-024-inflorance"  
/note="Vector: pCMVSPORT6; Site 1: SalI; Site 2: NotI;  
cDNA library from sugar beet, library provided by KWS  
Kleinwanzlebener Saatgut AG Bindeck, Germany; contact:  
b.schulz@kws.de; cloning sites SalI-NotI, primer sites and  
orientation:  
SP6-SalI-CCACGCGTCCG-5prime-cDNA-polyA-CC-NotI-T7;  
Sequencing granted in the context of the GABI-beet  
project, local PI: Dr. Katharina Schneider, coordinator:  
Prof. Christian Jung, Sequence submission managed by  
RZPD/GABI-Primary database: http://gabi.rzpd.de"

## ORIGIN

## Alignment Scores:

Pred. No.: 3.59e+04 Length: 16  
Score: 5.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.62% Indels: 0  
DB: 5 Gaps: 0

US-10-031-158-14 (1-58) x BQ583458 (1-16)

OY 8 Proleuphaphelenu 12

DB 15 CCCCTTTTCTTTT 1

## RESULT 3

AZ465954 19 bp DNA linear GSS 04-OCT-2000

LOCUS 1M0276816F Mouse 10kb plasmid U081M library Mus musculus genomic

DEFINITION clone U081M0276816 F, genomic survey sequence.

ACCESSION AZ465954

VERSION AZ465954.1 GI:10624079

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 19)

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,

JOURNAL Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,

MEDLINE Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von

COMMENT Niedermaier, A. and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: dunn@genetics.utah.edu

Insert Length: 1000 Std Error: 0.00

Plate: 0276 row: E column: 16

Seq primer: CATTGTAACGACGCGCAGT

Class: plasmid ends

# FEATURES High quality sequence stop: 19. Location/Qualifiers

1. .19

/organism="Mus musculus"

/mol\_type="genomic DNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="U08100276816"

/sex="Male"

/lab\_host="R. Coli strain XL10-Gold, T1-resistant, F-"

/clone\_1ib="Mouse 10kb plasmid U08100276816"

/note="Vector: PMD42nv, Purified genomic DNA from M. musculus C57BL/6J (male). Was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (G14732114[9d]/AF125072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent R. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

### Alignment Scores:

Pred. No.: 4.19e+04 Length: 19

Score: 5.00 Matches: 5

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 8.62% Indels: 0

DB: 8 Gaps: 0

US-10-031-158-14 (1-58) x AZ465954 (1-19)

Qy 8 Pheleuphephe 12

Db 2 CCCCTTTTCTTTT 16

## RESULT 4

AW246466 20 bp mRNA linear EST 07-JUN-2000

LOCUS 2821777.3pr1me NIH\_MGC\_7 Homo sapiens cDNA clone IMAGE:2821777 3',

DEFINITION mRNA sequence.

ACCESSION AW246466

VERSION AW246466.1 GI:6589459

KEYWORDS EST.

ORGANISM Homo sapiens (human)

SOURCE Homo sapiens

EuKaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 20)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Other ESTs: 2821777.5pr1me

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-r@mail.nih.gov

Tissue Procurement: DCTD/DP CNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML) DNA Sequencing by: Berkeley MGC sequencing project Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: www.bio.liml.gov/bhrp/image/image.html Base Calling / Quality Scores: PHRED from University of Washington Genome Center. Vector Trimming: cross\_match from University of Washington Genome Center

PHRAP suite. Poly-T identification: patmatch.pl from Berkeley Drosophila Genome Project. University of Washington Genome Center: http://www.genome.washington.edu Low Quality Sequence: 10 contiguous PHRED high quality bases following vector sequence. Very low Quality Sequence: Trace file contained 20 contiguous distinct peaks following vector sequence. Polyadenylation: Based upon the presence of a XhoI site followed by a run of 14 or more T residues at the beginning of the sequence, this cDNA insert was polyadenylated.

Plate: L10C7 row: L column: 2

High quality sequence stop: 10.

Location/Qualifiers

## FEATURES

### source

1. .20

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:2821777"

/tissue\_type="small cell carcinoma"

/cell\_line="MGC3"

/lab\_host="DH10B (phage-resistant)"

/clone\_1ib="NIH MGC 7"

/note="Organ: lung; Vector: pOT87; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

## Alignment Scores:

Pred. No.: 4.38e+04 Length: 20

Score: 5.00 Matches: 5

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 8.62% Indels: 0

DB: 2 Gaps: 0

US-10-031-158-14 (1-58) x AW246466 (1-20)

Qy 11 PheleuphepGlnLeu 15

Db 3 TTTTTCGACGCTT 17

## RESULT 5

CF282165 20 bp mRNA linear EST 14-AUG-2003

LOCUS 1487L--09-122.g1 Rice etiolated leaf plasmid cDNA library (1487L)

DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone 1487L--09-122, mRNA sequence.

ACCESSION CF282165

VERSION CF282165.1 GI:33659552

KEYWORDS EST.

ORGANISM Oryza sativa (japonica cultivar-group)

SOURCE Oryza sativa (japonica cultivar-group)

EuKaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 20)

Kim,J.S., Jun,K.M., Cheong,P.D., Kim,M.D., Lee,T.H., Shin,Y.C., Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.

Large-scale Sequencing Analysis of Rice ESTs

Unpublished (2003)

Contact: Nahm B.H.

Genomics and Genetics Institute, GreenGene Biotech Inc.; Division of BioScience and Bioinformatics, Myongji University

Yongin, Kyonggi, Korea

Tel: 82 31 320 6193

Fax: 82 31 321 6355

Email: bhnahm@bio.com, bhnahm@bio.myongji.ac.kr.

## FEATURES

### source

1. .20

Location/Qualifiers

/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nackdong"  
/db\_xref="taxon:39947"  
/clone="14RTL--09-122"  
/tissue\_type="leaf"  
/dev\_stage="14 days after germination"  
/lab\_host="E.coli DH10B"  
/clone\_id="Rice etiolated leaf plasmid cDNA library (14RTL)"  
/note="Vector: PCR4-TOPO. Site 1: EcoRI; mRNA was capped with oligoribonucleotides and then used as templates for RT-PCR."

## ORIGIN

## Alignment Scores:

Pred. No.:	4.38e+04	Length:	20
Score:	5.00	Matches:	5
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.62%	Indels:	0
DB:	7	Gaps:	0

US-10-031-158-14 (1-58) x CF282165 (1-20)

Qy 10 PhephepheleuGln 14

Db 4 TTTTTCCTTCTTA 18

## RESULT 6

AZ832043 20 bp DNA linear GSS 20-FEB-2001  
LOCUS 2M011211F Mouse 10kb plasmid UUGCM library Mus musculus genomic

DEFINITION clone UUGCM011211 F, genomic survey sequence.  
ACCESSION AZ832043  
VERSION AZ832043.1 GI:13001951

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Mus musculus (house mouse)  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 20)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,  
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von  
Niederhausern, A. and Wright, D., Weis, R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0112 row: 1 column: 11  
Seq primer: CGTTGTAAACGACGCGCCAGT  
Class: plasmid ends  
High quality sequence stop: 20.  
Location/Qualifiers

## FEATURES

## source

1..20  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGCM011211"  
/sex="Male"  
/lab\_host="E. coli strain XL10-Gold, TI-resistant, F-"  
/clone\_id="Mouse 10kb plasmid UUGCM library"  
/note="Vector: PWD42nv, Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adapted DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of PWD42 (g114732114[gblAF129072.1]), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adapted mouse DNA was annealed to  
adapted vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

## ORIGIN

Pred. No.:	4.38e+04	Length:	20
Score:	5.00	Matches:	5
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.62%	Indels:	0
DB:	8	Gaps:	0

US-10-031-158-14 (1-58) x AZ832043 (1-20)

Qy 9 Leuphepheleu 13

Db 2 CTATTTCTTCTTA 16

## RESULT 7

AZ957966 20 bp DNA linear GSS 27-APR-2001  
LOCUS 2M022510F Mouse 10kb plasmid UUGCM library Mus musculus genomic

DEFINITION clone UUGCM022510 F, genomic survey sequence.  
ACCESSION AZ957966  
VERSION AZ957966.1 GI:13829193

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Mus musculus (house mouse)  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 20)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,  
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von  
Niederhausern, A. and Wright, D., Weis, R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0225 row: 1 column: 01  
Seq primer: CGTTGTAAACGACGCGCCAGT  
Class: plasmid ends  
High quality sequence stop: 20.  
Location/Qualifiers

## FEATURES

## source

1..20  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"



/clone="UUGCM0225I01"  
/sex="female"  
/lab\_host="B. coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUGCM library"  
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptor complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

## Alignment Scores:

Pred. No.: 4.38e+04 Length: 20  
Score: 5.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.62% Indels: 0  
DB: 8 Gaps: 0

US-10-031-158-14 (1-58) x AZ957966 (1-20)

QY 8 Proleuphaphphe 12  
|||||  
6 CCCCTTTT TTTT 20

Db

RESULT 8 AG189044 20 bp DNA linear GSS 06-MAR-2004  
LOCUS Pan troglodytes DNA, clone: RP43-063F05.TU, genomic survey  
DEFINITION  
ACCESSION AG189044  
VERSION AG189044.1 GI:45221220  
KEYWORDS GSS.  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM Pan troglodytes

REFERENCE 1 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.  
2 (bases 1 to 20)

AUTHORS Park, H., Kim, Y., Kim, S., Han, Y., Woo, T., Park, K., Eun, C.J.,  
Hoon, S.T., Chu, M., Kim, H., Joo, S., Kim, C., Song, W. and Yoo, H.  
TITLE BAC end sequences of library RP-43  
JOURNAL Unpublished  
AUTHORS

REFERENCE 2 (bases 1 to 20)  
AUTHORS Park, H., Kim, Y., Kim, S., Han, Y., Woo, T., Park, K., Eun, C.J.,  
Hoon, S.T., Chu, M., Kim, H., Joo, S., Kim, C., Song, W. and Yoo, H.  
TITLE Direct Submission  
JOURNAL Submitted (07-JAN-2002) Hong-Seog Park, Korea Research Institute of  
AUTHORS BioScience and Biotechnology (KRIBB), Genome Research Center (GRC);  
52, Oun-dong, Yuseong-gu, Daejeon 305-333, Korea  
(E-mail: redstone@mail.krrib.re.kr, URL: http://phs.grc.krrib.re.kr/,  
Tel:82-42-866-7181, Fax:82-42-860-4403)  
Clones are derived from the chimpanzee BAC library RP-43 This BAC  
end was generated during the RAD process and may have higher chance  
of clone tracking errors.

## COMMENT

Sequencing: TJ

LIBRARY Vector : pBACE3.6  
R.Site 1 : EcorI

FEATURES  
source R.Site 2 : EcorI.  
Location/Qualifiers  
1..20

/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="RP43-063F05.TU"  
/sex="male"  
/cell\_type="lymphocytes"  
/clone\_lib="RP-43 Chimpanzee Male BAC Library"

## ORIGIN

## Alignment Scores:

Pred. No.: 4.38e+04 Length: 20  
Score: 5.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.62% Indels: 0  
DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x AG189044 (1-20)

QY 8 Proleuphaphphe 12  
|||||  
5 CCCCTTTT TTTT 19

Db

RESULT 9 CL660020 20 bp DNA linear GSS 09-JUL-2004  
LOCUS PRI0135d\_H08 - PRI0135d.B21 (20) Mixed stage fosmid library of P.  
DEFINITION pacificus var. California Pristionchus pacificus genomic, genomic  
survey sequence.  
ACCESSION CL660020  
VERSION CL660020.1 GI:50144803  
KEYWORDS GSS.  
SOURCE Pristionchus pacificus  
ORGANISM Pristionchus pacificus  
Bukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;  
Neodiplogasteridae; Pristionchus.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Sriniwasan, J., Otto, G.W., Kahlow, U., Geisler, R. and Sommer, R.J.  
TITLE AppDB: an AcceDB database for the nematode satellite organism  
JOURNAL Pristionchus pacificus  
Nucleic Acids Res. 32 (1), D421-D422 (2004)  
COMMENT Contact: Sommer RJ  
Evolutionary Biology  
Max-Planck-Institute for Developmental Biology  
Spemannstr. 37-39, Tuebingen D-72076, Germany  
Tel: 00497071601371  
Fax: 00497071601498  
Email: ralf.sommer@uebingen.mpg.de  
This library was generated at Caltech, Pasadena, USA and end  
sequenced at Vancouver, Canada.  
Seq primer: T7  
Classes: fosmid ends.

FEATURES  
source Location/Qualifiers  
1..20

/organism="Pristionchus pacificus"  
/mol\_type="genomic DNA"  
/strain="California"  
/db\_xref="taxon:54126"  
/clone\_lib="Mixed stage fosmid library of P. pacificus  
var. California"  
/note="Vector: pBplfos-5 Fosmid vector"

## ORIGIN

## Alignment Scores:

Pred. No.: 4.38e+04 Length: 20  
Score: 5.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.62% Indels: 0  
DB: 9 Gaps: 0



LOCUS	AZ3411108	21 bp	DNA	linear	GSS 29-SEP-2000
DEFINITION	1M0073014F Mouse 10kb plasmid UGCM1 library Mus musculus genomic clone UGCM0073014 F, genomic survey sequence.				
ACCESSION	AZ3411108				
VERSION	AZ3411108.1	GI:10417030			
KEYWORDS	GSS.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 21) Dunn,D., Aoyagi,A., Barber,M., Baccorn,T., Duval,B., Hamli,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,B., Pedersen,T., Reilly,M., Rose,R., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.				
TITLE	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Robert B. Weiss University of Utah Genome Center Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLCC, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert Length: 10000 Std Error: 0.00 Plate: 0073 row: 0 column: 14 Seq primer: CGTGTAAACGACGCGCCAGT Class: plasmid end High quality sequence stop: 21. Location/Qualifiers 1..21				
FEATURES					
SOURCE					

1. 21  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="VUGC1M0073014"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_1ib="Mouse 10kb plasmid VUGC1M library"  
/note="Vector: PWD24n; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adaptorized DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pMD24 (g[473214.19b]AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adaptorized mouse DNA was annealed to  
adaptorized vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

ORIGIN	
Alignment Scores:	
Pred. No.:	4.58e+04
Score:	5.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	8.62%
DB:	8
Gaps:	0
Length:	21
Matches:	5
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-10-031-158-14 (1-58) x AZ3411  
QY 8 Proleuphpephe 12

```

Db          19  CCCCCCTTTTCTTTT 5
RESULT 13
LOCUS      AZ589400.1
DEFINITION 21 bp DNA linear GSS 13-DEC-2000
ACCESSION  J0396C23F Mouse 10kb plasmid U06C1M library Mus musculus genomic
VERSION     AZ589400
KEYWORDS    AZ589400.1 GI:11711590
SOURCE      GSS.
ORGANISM    Mus musculus (house mouse)
REFERENCE   Mus musculus
AUTHORS     Bukaryotis, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 21)
            Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Ibrahim,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Relliy,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            Unpublished (2000)
            Contact: Robert B. Weiss
            University of Utah Genome Center
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0398 row: C column: 23
            Seq primer: CGTTGTAAACGACGCGCACT
            Class: plasmid
            High quality sequence stop: 21.
            Location/Qualifiers
            1..21
FEATURES
SOURCE

```

/organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC1M0398C23"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_1lb="Mouse 10kb plasmid UUGC1M library"  
 /note="Vector: PBD229; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-prepared with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 [gi|4732114|gb|AF12072.1], a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor-competent DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN :	
Alignment Scores:	
Pred. No.:	4.58e+04
Score:	5.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	8.62%
Length:	21
Matches:	0
Conservative:	0
Mismatches:	0
Indels:	0

Query: Match:

DB: 8 Gaps: 0

QY 8 ProLeuphPhephe 12  
 |||||  
 5 CCCCTTTTCTTTT 19

RESULT 14  
 AZ806282/c 21 bp DNA linear GSS 20-FEB-2001

LOCUS 2M0068A15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
 DEFINITION clone UUGC2M0068A15 F, genomic survey sequence.

ACCESSION AZ806282  
 VERSION AZ806282.1 GI:12967093

KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 21)  
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,  
 Rellily,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von  
 Niederhausern,A. and Wright,D., Weiss,R.  
 Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts

JOURNAL Unpublished (2000)  
 COMMENT Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 1000 Std Error: 0.00  
 Plate: 0068 row: A column: 15  
 Seq primer: CGTGTAAACGACGCGCAGT  
 Class: plasmid ends  
 High quality sequence stop: 21.

FEATURES  
 source

1. 21  
 Location/Qualifiers

/organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC2M0068A15"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_1lb="Mouse 10kb plasmid UUGC1M library"  
 /note="Vector: PMD42nv; Purified genomic DNA from M.  
 musculus C57BL/6J (male) was obtained from the Jackson  
 Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA  
 was hydrodynamically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adaptor DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of PMD42 (gi|4732114|gb|AF129072.1), a copy-number  
 inducible derivative of plasmid R1. The vector was ligated  
 with adaptors complementary to the insert adaptors and  
 purified. The sheared, adaptor mouse DNA was annealed to  
 adaptor vector DNA, and transformed into  
 chemically-competent E. coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

Pred. No.: 4.58e+04 Length: 21  
 Score: 5.00 Matches: 5  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 8.62% Indels: 0  
 DB: 8 Gaps: 0

US-10-031-158-14 (1-58) x AZ806282 (1-21)

QY 7 SerProLeuphPhephe 11  
 |||||  
 16 TCCCTCTTTCTTT 2

RESULT 15  
 AZ850337/c 21 bp DNA linear GSS 21-FEB-2001

LOCUS 2M0152H11F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
 DEFINITION clone UUGC2M0152H11 F, genomic survey sequence.

ACCESSION AZ850337  
 VERSION AZ850337.1 GI:13035243

KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 21)  
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,  
 Rellily,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von  
 Niederhausern,A. and Wright,D., Weiss,R.  
 Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts

JOURNAL Unpublished (2000)  
 COMMENT Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 1000 Std Error: 0.00  
 Plate: 0152 row: H column: 11  
 Seq primer: CGTGTAAACGACGCGCAGT  
 Class: plasmid ends  
 High quality sequence stop: 21.

FEATURES  
 source

1. 21  
 Location/Qualifiers

/organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC2M0152H11"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_1lb="Mouse 10kb plasmid UUGC1M library"  
 /note="Vector: PMD42nv; Purified genomic DNA from M.  
 musculus C57BL/6J (male) was obtained from the Jackson  
 Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA  
 was hydrodynamically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adaptor DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of PMD42 (gi|4732114|gb|AF129072.1), a copy-number  
 inducible derivative of plasmid R1. The vector was ligated  
 with adaptors complementary to the insert adaptors and  
 purified. The sheared, adaptor mouse DNA was annealed to  
 adaptor vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

ORIGIN

Alignment Scores:

Pred. No.:	4.58e+04	Length:	21
Score:	5.00	Matches:	5
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.62%	Indels:	0
DB:	8	Gaps:	0

US-10-031-158-14 (1-58) x AZ850337 (1-21)

Qy	6	ProSePProLeuphe	10
Db	16	CCCTCCCCCTCTCTTT	2

Search completed: October 16, 2005, 11:00:30  
Job time : 2985 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 16, 2005, 07:40:42 ; Search time 420 Seconds

(without alignments)  
817,488 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 58

Sequence: 1 MOMPSPPLFFFLQLKQSS.....RYGKKRATRFMDRRGTP 58

Scoring table: OLIGO  
Xgapop 60.0, Xgapext 60.0  
Ygapop 60.0, Ygapext 60.0  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3518947

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus.p2n.model -DEV=rlp  
-O=/cgn2.1/USPTO\_spool\_p/US10031158/runat\_14102005\_135842\_19595/app\_query.fasta\_1.199  
-DB=N\_Geneseq\_16Dec04 -QFMT=fastap -SUPFIX=p2n\_olig\_gz30.rng -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bases -START=1 -END=-1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MINLEN=30  
-USER=US10031158@cgn2.1 1.708 @runat\_14102005\_135842\_19595 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEOUTER -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOC  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_16Dec04:\*  
1: geneeqn1980s:\*  
2: geneeqn1990s:\*  
3: geneeqn2000s:\*  
4: geneeqn2001as:\*  
5: geneeqn2001bs:\*  
6: geneeqn2002as:\*  
7: geneeqn2002bs:\*  
8: geneeqn2003as:\*  
9: geneeqn2003bs:\*  
10: geneeqn2003cs:\*  
11: geneeqn2003ds:\*  
12: geneeqn2004as:\*  
13: geneeqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	8	13.8	27	4	AAFS6408 Human pro
C 2	8	13.8	27	4	AAFS6409 Human pro
C 3	7	12.1	21	2	AAQ04627 CTG prime
C 4	7	12.1	25	10	ADD28130 HIV-1 LTR
C 5	6	10.3	19	6	ABX97507 Human NOV

C 6	6	10.3	19	12	ADN62410	Adh62410 Human NOV
C 7	6	10.3	20	3	AAZ76949	Aaz76949 Human bia
C 8	6	10.3	20	4	AAH23198	Aah23198 Human MMT
C 9	6	10.3	20	8	ABZ10328	Abz10328 Hematopo
C 10	6	10.3	20	10	ADB54415	Adb54415 PCR prime
C 11	6	10.3	20	10	ADC69900	Adc69900 Primer 01
C 12	6	10.3	20	10	ADB77580	Adb77580 DRB3*0212
C 13	6	10.3	20	10	ABZ85409	Abz85409 Human o11
C 14	6	10.3	20	11	ABD21639	Abd21639 S100 calC
C 15	6	10.3	20	13	ADB89854	Adb89854 Human PCR
C 16	6	10.3	22	2	AAV15960	Aav15960 NBCCS (PT
C 17	6	10.3	23	6	AAJ39184	Aaj39184 Mouse 572
C 18	6	10.3	23	10	ADD28129	Add28129 HIV-1 LTR
C 19	6	10.3	25	9	ACI78562	Act78562 Human mic
C 20	6	10.3	25	9	ACK21141	Act21141 Human mic
C 21	6	10.3	25	9	ACI31211	Act31211 Human mic
C 22	6	10.3	25	9	ACI67768	Act67768 Human mic
C 23	6	10.3	25	9	ACI83555	Act83555 Human mic
C 24	6	10.3	25	9	ACI78563	Act78563 Human mic
C 25	6	10.3	25	10	ADB86664	Adb86664 Frizzled-
C 26	6	10.3	26	8	ABX94983	Abx94983 Barley RA
C 27	6	10.3	26	8	ABX94982	Abx94982 Barley RA
C 28	6	10.3	27	12	ADP70068	Adp70068 Novel mam
C 29	6	10.3	28	10	ADF43026	Adf43026 Bacterial
C 30	6	10.3	30	2	AAQ45800	Aaq45800 HBV ampli
C 31	6	10.3	30	2	AAV07797	Aav07797 HBV, 71 am
C 32	6	10.3	30	2	AAV83026	Aav83026 Amplifier
C 33	6	10.3	30	6	ABX69492	Abx69492 Novel Hel
C 34	6	10.3	30	10	ADC10250	Adc10250 Human NOV
C 35	6	10.3	30	12	ADM18572	Adm18572 Human sub
C 36	6	10.3	30	12	ADP08610	Adp08610 PCR prime
C 37	5	8.6	15	4	AAH46713	Aah46713 Type 11 p
C 38	5	8.6	15	4	ABA02603	Abao2603 PTEN targ
C 39	5	8.6	15	4	AAFS2874	Aafs2874 IGF-1 o11
C 40	5	8.6	15	4	AAFS2874	Aafs2874 IGF-1 o11
C 41	5	8.6	16	2	AAQ25455	Aaq25455 Purine ri
C 42	5	8.6	16	2	AAQ44151	Aaq44151 Sequence
C 43	5	8.6	16	2	AAQ43987	Aaq43987 HIV-1 tr1
C 44	5	8.6	16	2	AAQ45376	Aaq45376 Polypurin
C 45	5	8.6	16	2	AAQ95127	Aaq95127 Oligonuci

## ALIGNMENTS

RESULT 1	AAFS6408/c	AAFS6408 standard; DNA; 27 BP.
ID	AAFS6408	standard; DNA; 27 BP.
XX	XX	
AC	AAFS6408;	
XX	XX	
DT	12-APR-2001	(first entry)
XX	XX	
DE	Human prostate specific TCRgamma transcript PCR primer TCRgamma-R2.	
XX	XX	
KW	Human; TARP; prostate cancer; breast cancer; immunotherapy;	
KW	T cell receptor gamma alternate reading frame protein; TCRgamma;	
KW	PCR primer; ss.	
XX	XX	
OS	Homo sapiens.	
XX	XX	
PN	WO200104309-A1.	
XX	XX	
PD	18-JAN-2001.	
XX	XX	
PF	12-JUL-2000; 2000WO-US019039.	
XX	XX	
PR	13-JUL-1999; 99US-0143560P.	
XX	XX	
PR	01-OCT-1999; 99US-0157471P.	
XX	XX	
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
XX	XX	
PI	Pastan I, Essand M, Lee B, Vasmatazis G, Wolfgang C;	

DR WPI; 2001-081050/09.  
 XX Isolated T-cell receptor gamma alternate reading frame protein useful for  
 PT diagnosing and raising an immune response to prostate cancer and breast  
 PT cancer.  
 XX  
 PS Disclosure; Fig 8; 85pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of the  
 CC human T cell receptor alternate reading frame protein (TARP). This  
 CC protein is expressed in prostate and breast cancer cells at higher levels  
 CC than normal and so can be used in the immunotherapy of these cancers, as  
 CC well as their detection and prevention  
 XX  
 SQ Sequence 27 BP; 6 A; 4 C; 8 G; 9 T; 0 U; 0 Other;  
 XX  
 Alignment Scores:  
 Pred. No.: 12 Length: 27  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 13.79% Indels: 0  
 DB: 4 Gaps: 0  
 US-10-031-158-14 (1-58) x AAF56408 (1-27)  
 QY 13 LeuGINLeuLYsGInSeRser 20  
 DB 26 CTTCAATTGCTGAAACAAGCTCC 3  
 RESULT 2  
 AAF56409/C  
 ID AAF56409 standard; DNA; 27 BP.  
 XX  
 AC AAF56409;  
 XX  
 DT 12-APR-2001 (first entry)  
 DE Human prostate specific TCRgamma transcript PCR primer TCRgamma-R3.  
 XX  
 DE Human; TARP; prostate cancer; breast cancer; immunotherapy;  
 XX  
 KW T cell receptor gamma alternate reading frame protein; TCRgamma;  
 KW PCR primer; ss.  
 XX  
 OS Homo sapiens.  
 OS  
 PN MO200104309-A1.  
 PN  
 PD 18-JAN-2001.  
 PD  
 PF 12-JUL-2000; 2000MO-US019039.  
 PF  
 PR 13-JUL-1999; 99US-0143560P.  
 PR  
 PR 01-OCT-1999; 99US-0157471P.  
 PR  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 PI Pastan I, Essand M, Lee B, Vasmatazis G, Wolfgang C;  
 PI  
 DR WPI; 2001-081050/09.  
 DR  
 PT Isolated T-cell receptor gamma alternate reading frame protein useful for  
 PT diagnosing and raising an immune response to prostate cancer and breast  
 PT cancer.  
 PT  
 PS Disclosure; Fig 8; 85pp; English.  
 PS  
 CC The present invention provides the protein and coding sequences of the  
 CC human T cell receptor alternate reading frame protein (TARP). This  
 CC protein is expressed in prostate and breast cancer cells at higher levels  
 CC than normal and so can be used in the immunotherapy of these cancers, as  
 CC well as their detection and prevention  
 CC

SQ Sequence 27 BP; 10 A; 6 C; 6 G; 5 T; 0 U; 0 Other;  
 XX  
 Alignment Scores:  
 Pred. No.: 12 Length: 27  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 13.79% Indels: 0  
 DB: 4 Gaps: 0  
 US-10-031-158-14 (1-58) x AAF56409 (1-27)  
 QY 23 LeuGInHlsthPhenAlPhaleu 30  
 DB 26 CTGAACATACCTTGTCTCTTG 3  
 RESULT 3  
 AA004627/C  
 ID AA004627 standard; DNA; 21 BP.  
 XX  
 AC AA004627;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 31-OCT-2002 (revised)  
 DT 02-OCT-1990 (first entry)  
 DE CTG primer for T-cell receptor gene.  
 DE  
 KW Primer; T-cell receptor; C-region; gamma chain; lymphoid leukaemia;  
 KW lymphoma; PCR; ss.  
 XX  
 OS Synthetic.  
 OS  
 PN WO9004648-A.  
 PN  
 PD 03-MAY-1990.  
 PD  
 PF 20-OCT-1988; 88AU-00001057.  
 PF  
 PR 20-OCT-1988; 88AU-00001057.  
 PR  
 PA (MORL/) MORLEY A A.  
 PA  
 PI Bristol MJ;  
 PI  
 DR WPI; 1990-164030/21.  
 DR  
 PT Detection of lymphoid leukaemia and/or lymphoma - by determ. of  
 PT homogeneity or heterogeneity of the length of immunoglobulin and/or T-  
 PT receptor gene segments.  
 PT  
 PS Claim 12; Page 18; 23pp; English.  
 PS  
 CC The primer binds to the C-region of T-receptor gamma chain gene. Using  
 CC the primer T-receptor gene segments can then be used to determine homogeneity or  
 CC of the amplified fragments can then be used to determine homogeneity or  
 CC heterogeneity. This provides a rapid and sensitive diagnostic test for  
 CC leukaemia/lymphoma by detection of malignancy and determ. of the B- or T-  
 CC lymphocyte origin of the tumour. See also AA004603-004640. (Updated on 31  
 CC OCT-2002 to add missing OS field.) (Updated on 25-MAR-2003 to correct PA  
 CC field.)  
 CC  
 SQ Sequence 21 BP; 4 A; 4 C; 5 G; 8 T; 0 U; 0 Other;  
 XX  
 Alignment Scores:  
 Pred. No.: 86.1 Length: 21  
 Score: 7.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 12.07% Indels: 0  
 DB: 2 Gaps: 0  
 US-10-031-158-14 (1-58) x AA004627 (1-21)



QY 15 LeuLeuLyegInserSerArg 21  
 |||||  
 DB 21 TTGCTGAACAAAGCTCCAGA 1

RESULT 4  
 ID ADD28130  
 XX ADD28130 standard; DNA; 25 BP.  
 AC ADD28130;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE HIV-1 LTR region DNA fragment PNA1a.  
 XX  
 KM conjugate; peptide nucleic acid; PNA; transport mediator;  
 KM addressing protein; nuclear importation; penetratin family; tat gene;  
 KM rev gene; anti-HIV; primer; ss.  
 XX  
 OS Human immunodeficiency virus 1.  
 XX

Key Location/Qualifiers  
 FT misc\_feature 19..21  
 FT /\*tag= a  
 FT /note= "DNA sequence interrupted by an undisclosed  
 FT peptide linker"  
 FT 25  
 FT /\*tag= b  
 FT /mod\_base= OTHER  
 FT /note= "optionally pseudo-isocytosine or cytosine"  
 FT  
 PN WO2003006065-A2.  
 XX  
 PD 23-JAN-2003.  
 XX  
 PF 12-JUL-2002; 2002WO-DE002564.  
 XX  
 PR 12-JUL-2001; 2001DE-01033307.  
 XX  
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM STIFTUNG.  
 XX  
 PI Braun K, Waldeck W, Pipkorn R, Braun I, Debus J;  
 XX  
 DR WPI; 2003-256334/25.  
 XX  
 PT New conjugate for delivery to the cell nucleus, useful for treating  
 PT infection by human immune deficiency virus, comprises targeting peptides  
 PT and peptide nucleic acid.  
 XX  
 PS Disclosure; Page 8; 30pp; German.  
 XX

CC This invention describes a novel conjugate for mediating specific  
 CC transport to the cell nucleus of a peptide nucleic acid (PNA) that  
 CC hybridizes to an HIV (human immune deficiency virus) gene, or part of it.  
 CC The conjugate comprises a transport mediator for the cell membrane, an  
 CC addressing protein or peptide for nuclear importation and the PNA being  
 CC delivered. The transport mediator is a peptide or protein able to pass  
 CC through the plasma membrane, particularly derived from the penetratin  
 CC family protein. The PNA is designed to hybridize to at least part of the  
 CC tat or rev genes and is especially targeted to the long terminal repeat  
 CC or the polyurine tract, the central DNA flap, Nef, NCP7 or the gag slice  
 CC site, at the RNA level. The conjugate has the structure transport  
 CC mediator-addressing protein-PNA, particularly with a spacer (e.g.  
 CC polylysine or a synthetic polymer) between addressing protein and PNA.  
 CC The conjugate of the invention has anti-HIV activity. The transport  
 CC mediator/addressing protein combination provide rapid and efficient  
 CC transport of PNA to the nucleus and PNA is resistant to proteases and  
 CC nucleases. The conjugate can be used in both early and late stages of  
 CC infection. This sequence represents a fragment of HIV-1 LTR region used  
 CC to construct PNA's used in the method of the invention.  
 CC  
 XX  
 XX Sequence 25 BP; 2 A; 9 C; 0 G; 14 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 101  
 Score: 7.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 12.07%  
 DB: 10

US-10-031-158-14 (1-58) x ADD28130 (1-25)

QY 7 SerProLeuPhePheLeu 13  
 |||||  
 DB 4 TCCTCCCTTTCTTTCTTCT 24

RESULT 5  
 ID ABX97507/C  
 XX ABX97507 standard; DNA; 19 BP.  
 AC ABX97507;  
 XX  
 DT 20-MAY-2003 (first entry)  
 XX  
 DE Human NOV-associated forward primer from primer-probe set Ag3505.  
 XX  
 KM NOVX; cytosclastic; cardiac; antiarteriosclerotic; antiasthmatic; cancer;  
 KM hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;  
 KM human; PCR; primer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200272757-A2.  
 XX  
 PD 19-SEP-2002.  
 XX  
 PF 08-MAR-2002; 2002WO-US006908.  
 XX  
 PR 08-MAR-2001; 2001US-0274101P.  
 PR 08-MAR-2001; 2001US-0274194P.  
 PR 08-MAR-2001; 2001US-0274281P.  
 PR 08-MAR-2001; 2001US-0274322P.  
 PR 09-MAR-2001; 2001US-0274849P.  
 PR 12-MAR-2001; 2001US-0275235P.  
 PR 13-MAR-2001; 2001US-0275578P.  
 PR 13-MAR-2001; 2001US-0275579P.  
 PR 13-MAR-2001; 2001US-0275601P.  
 PR 14-MAR-2001; 2001US-0276000P.  
 PR 16-MAR-2001; 2001US-0276776P.  
 PR 19-MAR-2001; 2001US-0276994P.  
 PR 20-MAR-2001; 2001US-0277239P.  
 PR 20-MAR-2001; 2001US-0277321P.  
 PR 20-MAR-2001; 2001US-0277327P.  
 PR 21-MAR-2001; 2001US-0277791P.  
 PR 22-MAR-2001; 2001US-0277833P.  
 PR 23-MAR-2001; 2001US-0278152P.  
 PR 26-MAR-2001; 2001US-0278894P.  
 PR 27-MAR-2001; 2001US-0278999P.  
 PR 27-MAR-2001; 2001US-0279036P.  
 PR 28-MAR-2001; 2001US-0279344P.  
 PR 30-MAR-2001; 2001US-027738P.  
 PR 30-MAR-2001; 2001US-0279959P.  
 PR 30-MAR-2001; 2001US-0280233P.  
 PR 02-APR-2001; 2001US-0280802P.  
 PR 02-APR-2001; 2001US-0280822P.  
 PR 02-APR-2001; 2001US-0280900P.  
 PR 04-APR-2001; 2001US-0281194P.  
 PR 13-APR-2001; 2001US-0283675P.  
 PR 30-APR-2001; 2001US-0287424P.  
 PR 02-MAY-2001; 2001US-0288066P.  
 PR 03-MAY-2001; 2001US-0288342P.  
 PR 03-MAY-2001; 2001US-0288528P.  
 PR 15-MAY-2001; 2001US-0291190P.  
 PR 16-MAY-2001; 2001US-0291099P.  
 PR 16-MAY-2001; 2001US-0291240P.

PR 30-MAY-2001; 2001US-0294485P.  
 PR 31-MAY-2001; 2001US-0294889P.  
 PR 31-MAY-2001; 2001US-0294899P.  
 PR 18-JUN-2001; 2001US-0299027P.  
 PR 19-JUN-2001; 2001US-0299030P.  
 PR 19-JUN-2001; 2001US-0299310P.  
 PR 10-JUL-2001; 2001US-0304354P.  
 PR 31-JUL-2001; 2001US-0309198P.  
 PR 16-AUG-2001; 2001US-0312903P.  
 PR 12-SEP-2001; 2001US-0318462P.  
 PR 27-SEP-2001; 2001US-0325430P.  
 PR 27-SEP-2001; 2001US-0325681P.  
 PR 18-OCT-2001; 2001US-0330380P.  
 PR 31-OCT-2001; 2001US-0335301P.  
 PR 14-NOV-2001; 2001US-0332172P.  
 PR 14-NOV-2001; 2001US-0332272P.  
 PR 14-NOV-2001; 2001US-0333184P.  
 PR 21-NOV-2001; 2001US-0333272P.  
 PR 03-DEC-2001; 2001US-0337426P.  
 PR 03-DEC-2001; 2001US-0338092P.  
 PR 04-DEC-2001; 2001US-0337185P.  
 PR 03-JAN-2002; 2002US-0345705P.  
 PR 07-MAR-2002; 2002US-00092900.

## (CURA-) CURAGEN CORP.

PA Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CE, Li L,  
 PI Zernhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R,  
 PI Patnursajan M, Gangoli E, Vernet CM, Guo X, Tchiernev V,  
 PI Fernandes ER, Casman SJ, Malyanar UM, Gerlach V, Liu Y, Anderson D,  
 PI Spaderma SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP,  
 PI Lepley DM, Rieger DK,  
 DR WPI; 2002-723332/78.

XX NOVX polypeptides and polynucleotides, useful for preventing or treating  
 PT a disorder associated with aberrant NOVX expression or activity e.g.,  
 PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial  
 PT asthma.

XX Example C; Page 957; 1103pp; English.

XX This invention describes novel human NOVX polypeptides which have  
 CC cytostatic, cardiant, antiarteriosclerotic, antilasthmatic and hypotensive  
 CC activity. Pharmaceutical compositions comprising the NOVX proteins or  
 CC nucleic acid molecules or NOVX antibodies are useful for preventing or  
 CC treating a disorder associated with aberrant NOVX expression or activity  
 CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial  
 CC asthma. The products of the invention can be used for gene therapy or in  
 CC a vaccine. ABX13460-ABX13462 and ABX97186-ABX97593 represent PCR primers  
 CC and probes used in the amplification and isolation of the NOVX  
 CC polynucleotides represented in ABX97008-ABX97185 which encode the  
 CC polypeptides represented in ABX95041-ABX95218

XX Sequence 19 BP; 4 A; 6 C; 5 G; 4 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 714 Length: 19  
 Score: 6.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.34% Indels: 0  
 DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x ABX97507 (1-19)

OY 48 ThruArgPheTrpAspPro 53

DB 19 ACTAGATTCGGACCCG 2

RSULT 6  
 ADN62410/C  
 ID ADN62410 standard; DNA; 19 BP.  
 XX  
 AC ADN62410;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX  
 DE Human NOV98a RTQ-PCR forward primer.  
 XX  
 KW Human; ss; PCR; NOVX; diabetes; obesity; infectious disease; anorexia;  
 KW cancer-associated cachexia; cancer; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; immune disorder;  
 KW haematopoietic disorder; dyslipidaemia; chronic disease; primer; RTQ-PCR;  
 KW real time quantitative PCR.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2004043382-A1.  
 XX  
 PD 04-MAR-2004.  
 XX  
 PF 07-MAR-2002; 2002US-00092900.  
 XX  
 XX 08-MAR-2001; 2001US-0274191P.  
 XX 08-MAR-2001; 2001US-0274194P.  
 XX 08-MAR-2001; 2001US-0274281P.  
 XX 08-MAR-2001; 2001US-0274322P.  
 XX 09-MAR-2001; 2001US-0274849P.  
 XX 12-MAR-2001; 2001US-0275235P.  
 XX 13-MAR-2001; 2001US-0275578P.  
 XX 13-MAR-2001; 2001US-0275579P.  
 XX 13-MAR-2001; 2001US-0275601P.  
 XX 14-MAR-2001; 2001US-0276000P.  
 XX 16-MAR-2001; 2001US-0276767P.  
 XX 19-MAR-2001; 2001US-0276994P.  
 XX 20-MAR-2001; 2001US-0277239P.  
 XX 20-MAR-2001; 2001US-0277321P.  
 XX 20-MAR-2001; 2001US-0277327P.  
 XX 20-MAR-2001; 2001US-0277338P.  
 XX 21-MAR-2001; 2001US-0277791P.  
 XX 22-MAR-2001; 2001US-0277833P.  
 XX 23-MAR-2001; 2001US-0278152P.  
 XX 26-MAR-2001; 2001US-0278894P.  
 XX 27-MAR-2001; 2001US-0278999P.  
 XX 27-MAR-2001; 2001US-0279036P.  
 XX 28-MAR-2001; 2001US-0279344P.  
 XX 30-MAR-2001; 2001US-0279955P.  
 XX 30-MAR-2001; 2001US-0280233P.  
 XX 02-APR-2001; 2001US-0280802P.  
 XX 02-APR-2001; 2001US-0280822P.  
 XX 02-APR-2001; 2001US-0280900P.  
 XX 04-APR-2001; 2001US-0281444P.  
 XX 13-APR-2001; 2001US-0283675P.  
 XX 30-APR-2001; 2001US-0287424P.  
 XX 02-MAY-2001; 2001US-0288066P.  
 XX 03-MAY-2001; 2001US-0288342P.  
 XX 03-MAY-2001; 2001US-0288528P.  
 XX 15-MAY-2001; 2001US-0291190P.  
 XX 16-MAY-2001; 2001US-0291099P.  
 XX 16-MAY-2001; 2001US-0291240P.  
 XX 30-MAY-2001; 2001US-0294485P.  
 XX 31-MAY-2001; 2001US-0294889P.  
 XX 31-MAY-2001; 2001US-0294899P.  
 XX 18-JUN-2001; 2001US-0299027P.  
 XX 19-JUN-2001; 2001US-0299303P.  
 XX 19-JUN-2001; 2001US-0299310P.  
 XX 10-JUL-2001; 2001US-0304354P.  
 XX 31-JUL-2001; 2001US-0309198P.  
 XX 16-AUG-2001; 2001US-0312903P.  
 XX 10-SEP-2001; 2001US-0318462P.  
 XX 12-SEP-2001; 2001US-0318770P.  
 XX 27-SEP-2001; 2001US-0325430P.

27-SEP-2001; 2001US-0325681P.  
 PR 18-OCT-2001; 2001US-0330380P.  
 PR 31-OCT-2001; 2001US-0335301P.  
 PR 14-NOV-2001; 2001US-0332172P.  
 PR 14-NOV-2001; 2001US-0332271P.  
 PR 14-NOV-2001; 2001US-0332272P.  
 PR 14-NOV-2001; 2001US-0333184P.  
 PR 14-NOV-2001; 2001US-0333272P.  
 PR 21-NOV-2001; 2001US-0332094P.  
 PR 03-DEC-2001; 2001US-0337426P.  
 PR 03-DEC-2001; 2001US-0338092P.  
 PR 04-DEC-2001; 2001US-0337185P.  
 PR 03-JAN-2002; 2002US-0345705P.  
 XX  
 (PADI/) PADIGARU M.  
 PA (SPYT/) SPYTEK K. A.  
 PA (SHEN/) SHENOV S. G.  
 PA (TAUP/) TAUPIER R. J.  
 PA (PENNA/) PENNA C E A.  
 PA (LILL/) LI L.  
 PA (ZERR/) ZERRUSSEN B D.  
 PA (GUSEV/) GUSEV V Y.  
 PA (JIMW/) JI W.  
 PA (GORM/) GORMAN L.  
 PA (MILL/) MILLER C E.  
 PA (KEKU/) KEKUDA R.  
 PA (PATT/) PATURAJAN M.  
 PA (GANG/) GANGOLLI E A.  
 PA (VERN/) VERNET C A M.  
 PA (GUOX/) GUO X S.  
 PA (TCHER/) TCHERNIEV V T.  
 PA (FERN/) FERNANDES E R.  
 PA (CASW/) CASMAN S J.  
 PA (MALY/) MALYANKAR U M.  
 PA (GERL/) GERLACH V.  
 PA (LITU/) LITU Y.  
 PA (ANDE/) ANDERSON D W.  
 PA (SPAD/) SPADERNA S K.  
 PA (CATT/) CATTERTON E.  
 PA (LEIT/) LEITE M W.  
 PA (ZHON/) ZHONG H.  
 PA (ALSO/) ALSOOROCK J P.  
 PA (LEPL/) LEPLEY D M.  
 PA (RIEG/) RIEGER D K.  
 PA (BURG/) BURGESS C E.  
 XX  
 Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CE, Li L,  
 PI Zernusen BD, Gusev VY, Ji W, Gorman L, Miller CE, Kekuda R,  
 PI Paturajan M, Gangolli EA, Vernet CM, Guo XS, Tcherniey VT,  
 PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y,  
 PI Anderson DM, Spaderna SK, Catterton E, Leite MW, Zhong H;  
 PI Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;  
 XX  
 WPI; 2004-225693/21.  
 XX  
 New NOVX polypeptides and nucleic acid molecules useful for diagnosing,  
 PT preventing or treating NOVX-associated disorders, e.g. cancer, diabetes,  
 PT infection or obesity, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 XX  
 Example C; SEQ ID NO 680; 786bp; English.  
 XX  
 The invention relates to an isolated polypeptide (designated NOVX, or  
 CC NOVX-NOV127) comprising a sequence selected from 178 fully defined amino  
 CC acid sequences (and their mature forms, variants and fragments). Also  
 CC included are an isolated nucleic acid molecule encoding NOVX, a vector  
 CC comprising the nucleic acid, a cell comprising the vector, methods for  
 CC determining the presence or amount of the polypeptide or the nucleic acid  
 CC molecule in a sample, methods for determining the presence of or  
 CC predisposition to a disease associated with altered levels of expression  
 CC of the above polypeptide or nucleic acid molecule in a first mammalian  
 CC subject, a method for identifying an agent that binds to the above  
 CC polypeptide, a method for identifying a potential therapeutic agent for

CC use in the treatment of a pathology that is related to aberrant  
 CC expression or physiological interactions of the polypeptide, a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide and a method for modulating  
 CC the activity of the polypeptide cited above. The composition and methods  
 CC are useful for diagnosing, preventing or treating diseases such as  
 CC diabetes, obesity, infectious diseases, anorexia, cancer-associated  
 CC cachexia, cancer, neurodegenerative disorders like Alzheimer's disease or  
 CC Parkinson's disease, immune disorders, haematopoietic disorders,  
 CC dyslipidaemias, and other chronic diseases. These may also be used in  
 CC chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The polypeptides are also useful as vaccines. The  
 CC present sequence is an RT0-PCR (real time quantitative PCR) primer used  
 CC to assay tissue specific expression of a NOVX mRNA.  
 XX  
 SO Sequence 19 BP; 4 A; 6 C; 5 G; 4 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. NO.: 714 Length: 19  
 Score: 6.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.34% Indels: 0  
 DB: 12 Gaps: 0  
 US-10-031-158-14 (1-58) x ADM62410 (1-19)  
 QY 48 ThArghpethrPaSPPro 53  
 DB 19 ACTAGATTCTGGACCCG 2  
 RESULT 7  
 AAZ76949  
 ID AAZ76949 standard; DNA; 20 BP.  
 AC AAZ76949;  
 XX  
 DT 10-SEP-2001 (first entry)  
 XX  
 DE Human biallelic marker downstream amplification primer SEQ ID NO:11305.  
 XX  
 KW Human genome; biallelic marker; high density disequilibrium map;  
 KW genomic map; haplotype; phenotype; polymorphic base; genotyping;  
 KW haplotyping; hybridisation; identification; characterisation;  
 KW amplification; single nucleotide polymorphism; SNP; PCR primer;  
 KW diagnosis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO954500-A2.  
 XX  
 PD 28-OCT-1999.  
 XX  
 PF 21-APR-1999; 99WO-IB000822.  
 XX  
 PR 21-APR-1998; 98US-0082614P.  
 PR 23-NOV-1998; 98US-0109732P.  
 XX  
 PA (GENT ) GENSET.  
 XX  
 PI Cohen D, Blumenfeld M, Chumakov I;  
 XX  
 DR WPI; 2000-013267/01.  
 XX  
 PT Novel biallelic markers used to construct a high density disequilibrium  
 PT map of the human genome.  
 XX  
 PS Claim 9; Page 2640; 2745bp; English.  
 XX  
 CC AAZ65654 to AAZ69578 represent human biallelic markers from the present  
 CC invention, which contain a polymorphic base at position 24 of their  
 CC nucleotide sequences. AAZ69579 to AAZ77440 represent amplification  
 CC primers for the biallelic markers. The biallelic markers of the invention

CC have a variety of uses: they can be used for high density mapping of the  
CC human genome, and in complex association studies and haplotyping studies.  
CC which are useful in determining the genetic basis for disease states.  
CC Compositions and methods of the invention can also be useful for the  
CC identification of the targets for the development of pharmaceutical  
CC agents and diagnostic methods, as well as the characterisation of the  
CC differential efficacious responses to and side effects from  
CC pharmaceutical agents acting on a disease as well as other treatment  
CC N.B. The SEQ ID NOS 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297 and  
CC 3367, are not actually given a sequence in the Sequence Listing from the  
CC present invention

XX  
XX  
SQ     Sequence 20 BP; 2 A; 5 C; 2 G; 11 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.:               749               Length:               20  
Score:                  6.00              Matches:              6  
Percent Similarity:    100.00%          Conservative:        0  
Best Local Similarity: 100.00%          Mismatches:         0  
Query Match:           10.34%          Indels:              0  
DB:                     3               Gaps:                0

US-10-031-158-14 (1-58) x AA276949 (1-20)

Oy                       8 ProLeuPhePhePheIeu 13  
Db                       3 CCCCTATTTTCTTG 20  
AAH23198  
ID     AAH23198 standard; DNA; 20 BP.  
XX  
XX     AAH23198;  
XX  
DT     17-SRP-2001 (first entry)  
DE     Human MMIF mRNA inhibiting antisense oligo ISIS #112694.

XX  
XX     Macrophage migration inhibitory factor; MMIF; antisense; neurological;  
KM     hyperproliferation; nootropic; antihormonal; immunosuppressive; human;  
KM     antiinflammatory; cyostatic; ss.  
OS  
OS     Synthetic.  
OS     Homo sapiens.  
FN     WO2001S3317-A1.  
PD  
PD     26-JUL-2001.  
PF  
PF     16-JAN-2001; 2001MO-US001475.  
PR  
PR     20-JAN-2000; 2000US-00489869.  
XX  
XX     (ISIS-) ISIS PHARM INC.  
PA  
PI     Murray SF, Cowsett LM, Wyatt JR;  
PI     WPT, 2001-451899/48.  
DR  
XX  
XX     New antisense compound(s) are useful to inhibit a nucleic acid molecule  
PT     encoding macrophage migration inhibitory factor.  
PS  
PS     Claim 3; Page 82; 105pp; English.

The invention relates to antisense oligonucleotides 8-30 nucleotides in  
length targeted to a nucleic acid molecule encoding macrophage migration  
inhibitory factor (MMIF), where the antisense compound specifically  
hybridizes with and inhibits the expression of MMIF. The antisense  
nucleotides are useful for the treatment of a disease or condition  
associated with MMIF such as neurological, hormonal, immune, inflammatory,  
or hyperproliferative disorder. Sequences AAH23191-268 represent chimeric  
MMIF mRNA expression

XX	SQ	Sequence	20 BP; 6 A; 7 C; 7 G; 0 T; 0 U; 0 Other;
XX		Alignment Scores:	
XX	Pred. No.:	749	Length: 20
XX	Score:	6.00	Matches: 6
XX	Percent Similarity:	100.00%	Conservative: 0
XX	Best Local Similarity:	100.00%	Mismatches: 0
XX	Query Match:	10.34%	Indels: 0
XX	DB:	4	Gaps: 0
XX	US-10-031-158-14 (1-58) x AAH23198 (1-20)		
XX	OY	52 AspproAARGGLyThr 57	
XX			
XX	Db	3 GACCGGGCAGAGGCACA 20	
XX	RESULT 9		
XX	ABZ10328/c		
XX	ID	ABZ10328 standard; DNA; 20 BP.	
XX	XX		
XX	AC	ABZ10328;	
XX	DT	16-JAN-2003 (first entry)	
XX	DE	Hematopoietic cell proliferation disorder related primer SEQ ID NO:468.	
XX	KW	Human; hematopoietic cell proliferation disorder; cytostatic;	
XX	KV	gene therapy; lymphocytic leukemia; acute myelogenous leukaemia;	
XX	KX	Cytosine methylation state; probe; primer; ss.	
XX	OS	Homo sapiens.	
XX	OS	Synthetic.	
XX	PN	WO200277272-A2.	
XX	PD	03-OCT-2002.	
XX	PF	26-MAR-2002; 2002MO-EP003401.	
XX	PR	26-MAR-2001; 2001US-0278333P.	
XX	PA	(BPG-) EPIDENOMICS AG.	
XX	P1	Berlin K, Braun A, Dietler J, Guetig D, Howe A, Mueller J;	
XX	P1	Lewin A, Piepenbrock C, Adorjan P, Grabs G, Jesche R, Leu B,	
XX	P1	Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;	
XX	P1	Schoepe I, Ziebarth H;	
XX	DR	WPI; 2003-018942/OI.	
XX	PT	Detecting and differentiating between hematopoietic cell proliferative	
XX	PT	disorders, comprises contacting a target nucleic acid with a reagent that	
XX	PS	distinguishes between methylated and non-methylated CpG dinucleotides.	
XX	PS	Claim 11; Page 36; 117pp; English.	
CC	The present invention describes a method for detecting and		
CC	differentiating between haematopoietic cell proliferative disorders		
CC	associated with at least 1 gene and/or their regulatory regions in a		
CC	subject. The method comprises contacting a target nucleic acid in a		
CC	biological sample obtained from the subject with at least 1 reagent,		
CC	which distinguishes between methylated and non-methylated CpG		
CC	dinucleotides within the target nucleic acid. ABZ0961 to ABZ1118		
CC	represent specifically claimed nucleotide sequences from the present		
CC	invention. Oligonucleotides from the present invention can be used for		
CC	differentiating between healthy hematopoietic cells and proliferative		
CC	disorder hematopoietic cells; for differentiating between acute		
CC	lymphocytic leukaemia and acute myelogenous leukaemia; as probes for		
CC	determining the cytosine methylation state and/or single nucleotide		
CC	polymorphisms (SNPs) of haematopoietic cell proliferation disorder		
CC	related sequences and their complements, and as primers for the		
CC	amplification of haematopoietic cell proliferation disorder related DNA		

CC sequences. The nucleotide sequences from the present invention can also  
CC be used for detecting a predisposition to, differentiation between  
CC subclones, diagnosis, prognosis, treatment and/or monitoring of  
CC haematopoietic cell proliferative disorders. The present method enables a  
CC highly specific classification of haematopoietic cell proliferative  
CC disorders allowing for improved and informed treatment of patients  
XX

SO Sequence 20 BP; 8 A; 0 C; 11 G; 1 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	749	Length:	20
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	8	Gaps:	0

US-10-031-158-14 (1-58) x ABZ10328 (1-20)

QY 5 ProProSeerProleuphe 10

ID ADB54415/C

DB 18 CCTCCTCTCCTCTCTTC 1

ADBS4415; standard; DNA; 20 BP.

ADBS4415;

04-DEC-2003 (first entry)

PCR primer 83 used to amplify genomic DNA region.

colon cell proliferative disorder; non methylated CpG dinucleotide;

cytostatic; cancer; adenoma; carcinoma; cytosine methylation state; ss;

PCR; primer.

Unidentified.

WO2003072821-A2.

04-SEP-2003.

27-FEB-2003; 2003WO-EP002035.

27-FEB-2002; 2002EP-00004551.

(EPIG-) EPIGENOMICS AG.

Adorjan P, Burger M, Maier S, Nimmrich I, Becker E, Lesche R;

Rujan T, Schmitt A;

WPI; 2003-731620/69.

Detecting and differentiating between colon cell proliferative disorders

associated with a gene or its regulatory regions comprises contacting a

target nucleic acid in a biological sample obtained from the subject with

a reagent.

Claim 15; Page 24; 74pp; English.

The invention relates to a novel method for detecting and differentiating  
between colon cell proliferative disorders associated with at least one  
gene or its regulatory regions. The method comprises contacting a target  
nucleic acid in a biological sample obtained from the subject with at  
least one reagent or a series of reagents, where the reagent or series of  
reagents, distinguishes between methylated and non methylated CpG  
dinucleotides within the target nucleic acid. The molecules of the  
invention demonstrate cytosine activity whilst the method may useful  
for detecting and differentiating between colon cell proliferative  
disorders, including cancers such as colon adenoma and colon carcinoma.  
The DNA (peptide nucleic acid)-oligomers are useful as probes for  
determining cytosine methylation state or single nucleotide

CC polymorphisms. The current sequence is that of the PCR primer of the  
CC invention which was used to amplify the genomic DNA region.

SO Sequence 20 BP; 8 A; 0 C; 11 G; 1 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	749	Length:	20
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	10	Gaps:	0

US-10-031-158-14 (1-58) x ADB54415 (1-20)

QY 5 ProProSeerProleuphe 10

ID ADC69900/C

DB 18 CCTCCTCTCCTCTCTTC 1

ADC69900; standard; DNA; 20 BP.

ADC69900;

18-DEC-2003 (first entry)

Primer oligo used to amplify pretreated genomic DNA (SeqID 389).

PCR; primer; ss; lung cell proliferative disorder; CpG dinucleotide;

adenocarcinoma; squamous cell carcinoma; cytostatic; probe; PNA-oligomer;

cytosine methylation state.

Unidentified.

WO2003052135-A2.

26-JUN-2003.

10-DEC-2002; 2002WO-EP014026.

14-DEC-2001; 2001DE-01061625.

(EPIG-) EPIGENOMICS AG.

Burger M, Field JK, Genc B, Liloglou T, Lipacher E, Maier S;

Nimmrich I;

WPI; 2003-533029/50.

Detecting and differentiating cytosine methylation state of genomic DNA,

useful for diagnosing, treating prognosticating and/or monitoring lung

cell proliferative disorders e.g. adenocarcinoma and squamous cell

carcinoma.

Claim 11; SEQ ID NO 389; 58pp; English.

This invention relates to a novel method for detecting and  
differentiating between lung cell proliferative disorders associated with  
at least one gene and/or their regulatory regions. Specifically, it  
refers to a method comprising contacting a target nucleic acid in a  
biological sample with at least one reagent, wherein the reagent is able  
to distinguish between methylated and non-methylated CpG dinucleotides  
present in the target DNA. As such, it is possible to further  
differentiate and diagnose medical conditions including adenocarcinoma  
and squamous cell carcinoma, and their respective adjacent lung tissue.  
The present invention describes cytosine oligomers and PNA-oligomers  
CC that are useful as probes for determining the cytosine methylation state  
or single nucleotide polymorphisms (SNPs) of the target sequence. This  
oligonucleotide sequence is a primer oligomer used for the amplification  
of pretreated DNA (i.e. where unmethylated cytosine bases are converted  
to uracil), used in an exemplification of the invention.

SQ Sequence 20 BP; 8 A; 0 C; 11 G; 1 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 749 Length: 20  
 Score: 6.00 Matches: 6  
 Percent Similarity: 100.00% Conservatve: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.34% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-031-158-14 (1-58) x ADC69900 (1-20)  
 QY 5 ProProSerProLeuphe 10  
 DB 18 CCTCCCTCTCCTCTCTC 1  
 RESULT 12  
 ADE77580  
 ID ADE77580 standard; DNA; 20 BP.  
 AC ADE77580;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE DRB3\*0212 probe designed to analyse the HLA-DRB polymorphic region.  
 XX  
 XX HLA-DRB; probe; ss; human; multiplexed elongation assay;  
 XX enzymatic recognition;  
 KW cystic fibrosis conductance transmembrane regulator; CFTR;  
 KW human leukocyte antigen; HLA; genetic testing; carrier screening;  
 KW genotyping; profiling; polymorphic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003034029-A2.  
 PD 24-APR-2003.  
 XX  
 PF 15-OCT-2002; 2002WO-US033012.  
 XX  
 PR 15-OCT-2001; 2001US-0329427B.  
 PR 15-OCT-2001; 2001US-0329428B.  
 PR 15-OCT-2001; 2001US-0329619P.  
 PR 15-OCT-2001; 2001US-0329620P.  
 PR 14-MAR-2002; 2002US-0364416P.  
 XX  
 PA (BIOA-) BIOARRAY SOLUTIONS LTD.  
 PI Li AX, Hashmi G, Seul M;  
 DR WPI; 2003-393553/37.  
 PT Concurrent interrogation of a number of polymorphic sites, useful for  
 PT genetic testing, carrier screening, genetic profiling, and identity  
 PT testing, comprises conducting a multiplexed elongation assay using  
 PT probes.  
 PS Example 2; Page 38; 143pp; English.  
 XX  
 CC This invention relates to a novel method for the concurrent interrogation  
 CC of a number of polymorphic sites in the presence of, and without  
 CC interference from, non-designated polymorphic sites. Specifically, it  
 CC comprises conducting a multiplexed elongation assay by applying one or  
 CC more temperature cycles to achieve linear amplification of the target or  
 CC a combination of annealing and elongation steps under temperature-  
 CC controlled conditions. Furthermore, this detection method uses probe  
 CC extension or elongation and relies on enzymatic recognition, a superior  
 CC technique that no longer depends on differential hybridisation. The  
 CC present invention describes probes and methods useful for identifying or  
 CC detecting polymorphisms at one or more designated sites, such that they  
 CC can identify mutations within the cystic fibrosis conductance  
 CC transmembrane regulator (CFTR) or the human leukocyte antigen (HLA)  
 CC genes. In addition, concurrent interrogation of a multiplicity of

CC polymorphic sites is useful for genetic testing, carrier screening,  
 CC genotyping or genetic profiling, and identity testing. This  
 CC oligonucleotide is a probe designed to analyse the polymorphic region of  
 CC the HLA-DRB gene of the invention.  
 SQ Sequence 20 BP; 3 A; 5 C; 4 G; 8 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 749 Length: 20  
 Score: 6.00 Matches: 6  
 Percent Similarity: 100.00% Conservatve: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.34% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-031-158-14 (1-58) x ADE77580 (1-20)  
 QY 12 PhleuGlnLeuLeuLe 17  
 DB 1 TTCTTGACGCTGCTTAAG 18  
 RESULT 13  
 AB285409  
 ID AB285409 standard; DNA; 20 BP.  
 AC AB285409;  
 XX  
 DT 17-OCT-2003 (first entry)  
 XX  
 DE Human oligonucleotide sequence.  
 XX  
 XX Human; antisense; lung dysfunction; nasal airway dysfunction;  
 KW antiinflammatory steroid; ubiquinone; antiinflammatory; antiallergic;  
 KW antiaesthetic; hypotensive; immunosuppressive; cytostatic; gene therapy;  
 KW antisense gene therapy; respiratory; lung; adenosine sensitivity;  
 KW adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;  
 KW lung inflammation; respiratory disease; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200285308-A2.  
 PD 31-OCT-2002.  
 XX  
 PF 23-APR-2002; 2002WO-US013135.  
 XX  
 PR 24-APR-2001; 2001US-0286137P.  
 XX  
 PA (EPIC-) EPICGENESIS PHARM INC.  
 PI Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;  
 PI Miller S, Tang L, Shahabuddin S;  
 DR WPI; 2003-229219/22.  
 PT Pharmaceutical composition for treating ailments associated with impaired  
 PT respiration, has oligo(s) antisense to specific gene(s) or its  
 PT corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or  
 PT ubiquinone.  
 PS Claim 15; SEQ ID NO 651; 872pp; English.  
 XX  
 CC The invention relates to a novel pharmaceutical composition, which has a  
 CC first active agent comprising an oligonucleotide antisense to the  
 CC initiation codon, coding region, 5' or 3' end genomic flanking regions,  
 CC 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of  
 CC junctions of genes encoding a polypeptide associated with lung and/or  
 CC nasal airway dysfunction and a second active agent comprising an  
 CC antiinflammatory steroid and ubiquinone. A composition of the invention  
 CC has antiinflammatory, antiallergic, antiaesthetic, hypotensive,  
 CC immunosuppressive, and cytostatic activity. The composition may have a  
 CC use in antisense gene therapy. The composition is useful for treating or  
 CC preventing a respiratory, lung or malignant disease or condition, also

CC for enhancing the prophylactic or therapeutic respiratory effect of an  
CC antiinflammatory steroid in a subject, for reducing or depleting levels  
CC of, or reducing sensitivity to adenosine, reducing levels of adenosine  
CC receptor, producing bronchodilation, increasing levels of ubiquinone or  
CC lung surfactant in a subject's tissue, or treating bronchoconstriction,  
CC lung inflammation, lung allergies, or a respiratory disease or condition.  
CC Note: The sequence data for this patent is not represented in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 20 BP; 2 A; 13 C; 1 G; 4 T; 0 U; 0 Other;

XX Alignment Scores:

XX Pred. No.: 749 Length: 20  
XX Score: 6.00 Matches: 6  
XX Percent Similarity: 100.00% Conservative: 0  
XX Best Local Similarity: 100.00% Mismatches: 0  
XX Query Match: 10.34% Indels: 0  
XX DB: 10 Gaps: 0

US-10-031-158-14 (1-58) x ABZ85409 (1-20)

QY 4 PheProPserProleu 9  
DB 3 TTCCCTCCAGCCCTC 20

RESULT 14

ABD21639  
ID ABD21639 standard; DNA; 20 BP.

XX ABD21639;  
XX 29-JUL-2004 (first entry)

DE S100 calcium binding protein A2-derived oligo SEQ ID 651.

XX Human; antisense; bronchoconstriction; allergy; hyposecretion; pain;  
XX respiratory tract inflammation; adenosine sensitivity; lung; cancer;  
XX surfactant depletion; antiallergic; antiinflammatory; antiasthmatic;  
XX analgesic; hypotensive; immunosuppressive; cytostatic; cystic fibrosis;  
XX beta-adrenergic agonist; respiratory disease; pulmonary vasoconstriction;  
XX respiratory distress syndrome; allergic rhinitis; pulmonary hypertension;  
XX emphysema; chronic obstructive pulmonary disease; cancer; bronchitis;  
XX pulmonary transplantation rejection; ss; primer.

OS Homo sapiens.

XX WO200285309-A2.

XX 31-OCT-2002.

XX 23-APR-2002; 2002WO-US013143.

XX 24-APR-2001; 2001US-0286036P.

XX (EPIG-) EPIGENESIS PHARM INC.

XX NYce JM, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D,  
XX Miller S, Tang L, Shahabuddin S;

XX WPI; 2003-093058/08.

XX Pharmaceutical composition for treating asthma, has antisense  
XX oligonucleotide containing less percentage of adenosine, targeted to  
XX nucleic acids associated with lung airway or lung dysfunction, and  
XX bronchodilating agent.

XX Claim 15; SEQ ID NO 651; 763bp; English.

XX This invention describes a novel composition (a) a first active agent,  
XX comprising oligonucleotides, effective for alleviating  
XX bronchoconstriction, respiratory tract inflammation, allergies and  
XX reducing adenosine sensitivity, levels of adenosine (A) or (A) receptors,

CC surfactant depletion or hyposecretion, when administered to a mammal. The  
CC oligonucleotides are derived from a gene encoding or regulating  
CC expression of a target polypeptide associated with lung airway or lung  
CC dysfunction or cancer and can be anti-sense to the corresponding mRNA.  
CC The invention also describes a kit, that comprises: (a) a delivery  
CC device, in separate containers, (b) the oligonucleotides, (c)  
CC instructions for adding a carrier and for use of the kit. The composition  
CC of the invention has antiallergic, antiinflammatory, antiasthmatic,  
CC analgesic, hypotensive, immunosuppressive and cytostatic activity, is a  
CC beta-adrenergic agonist. The composition is useful for preventing or  
CC treating a respiratory, lung or malignant disease. The administered  
CC composition comprises oligo and is administered to reduce the production  
CC or availability, or to increase the degradation of the target mRNA or to  
CC reduce the amount of target polypeptide present in the lungs. The  
CC pulmonary obstruction, and/or bronchoconstriction and/or lung  
CC inflammation, allergies and/or surfactant hypoproduction are associated  
CC with a disease or condition such as pulmonary vasoconstriction,  
CC inflammation, allergies, asthma, impaired respiration, respiratory  
CC distress syndrome, emphysema, chronic obstructive pulmonary disease, pulmonary  
CC hypertension, emphysema, chronic obstructive pulmonary disease, pulmonary  
CC transplantation rejection, pulmonary infections, bronchitis or cancer.  
CC The reduced adenosine content of the anti-sense oligos corresponding to  
CC thymidines present in the target RNA serves to prevent the breakdown of  
CC the oligonucleotides into products that free adenosine into the system  
CC e.g., lung, brain, heart, kidney, etc, tissue environment and thereby, to  
CC prevent any unwanted effects due to it

XX Sequence 20 BP; 2 A; 13 C; 1 G; 4 T; 0 U; 0 Other;

XX Alignment Scores:

XX Pred. No.: 749 Length: 20  
XX Score: 6.00 Matches: 6  
XX Percent Similarity: 100.00% Conservative: 0  
XX Best Local Similarity: 100.00% Mismatches: 0  
XX Query Match: 10.34% Indels: 0  
XX DB: 11 Gaps: 0

US-10-031-158-14 (1-58) x ABD21639 (1-20)

QY 4 PheProPserProleu 9  
DB 3 TTCCCTCCAGCCCTC 20

RESULT 15

ADS89854/C  
ID ADS89854 standard; DNA; 20 BP.

XX ADS89854;

XX 18-NOV-2004 (first entry)

XX Human PCR primer SEQ ID NO:870.

XX ss; PCR; primer; human; cell proliferative disorder; breast; methylation;  
XX cytostatic; gene therapy; single nucleotide polymorphism; SNP.

OS Homo sapiens.

XX WO2004035803-A2.

XX 29-APR-2004.

XX 01-OCT-2003; 2003WO-EP010881.

XX 01-OCT-2002; 2002DE-01045779.

XX 07-JAN-2003; 2003DE-01000096.

XX 17-APR-2003; 2003DE-01017955.

XX (EPIG-) EPIGENOMICS AG.

XX Foekens J, Harbeck N, Koenig T, Maier S, Martens J, Model F,  
XX Nimrich I, Rujan T, Schmitt A, Schmitt M, Look MP, Marx A;

DR MPI: 2004-348468/32.

XX Predicting responsiveness of a subject with breast cell proliferative  
 PT disorder, useful for treating or differentiating breast cell  
 PT proliferative disorders comprises analyzing methylation pattern of a  
 PT genomic DNA from the subject.

XX Example 2; SEQ ID NO 870; 104bp; English.

CC The invention relates to a novel method for predicting the responsiveness  
 CC of a subject with a cell proliferative disorder of the breast tissues to  
 CC a therapy comprising analysing the methylation pattern of a target  
 CC nucleic acid by contacting at least one of the target nucleic acids in a  
 CC biological sample obtained from the subject prior to or during treatment.  
 CC The method of the invention has cytosolic activity, and may have a use  
 CC in gene therapy. The set of oligonucleotides comprising at least two of  
 CC the oligomers are useful for detecting the cytosine methylation state  
 CC and/or single nucleotide polymorphisms (SNPs) within the sequences. The  
 CC methods, nucleic acid, oligonucleotide, and kit are useful for the  
 CC treatment, characterisation, classification and/or differentiation, of  
 CC breast cell proliferative disorders. The method is also useful for  
 CC predicting the responsiveness of a subject with a cell proliferative  
 CC disorder of the breast tissues to a therapy. The present sequence is used  
 CC in the exemplification of the invention.

XX SQ Sequence 20 BP; 8 A, 0 C; 11 G; 1 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	749	Length:	20
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	13	Gaps:	0

US-10-031-158-14 (1-58) x ADS89854 (1-20)

QY 5 ProPterProleuphe 10

DB 18 CCTCCCTCTCTCTCTC 1

Search completed: October 16, 2005, 09:40:49  
 Job time : 423 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 16, 2005, 07:43:47 / Search time 1775 Seconds  
(without alignments)  
1583.325 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 58  
Sequence: 1 MCMFPSPPLFFLLQLKQSS.....RYIGKKRRATRWDPGRGTP 58

Scoring table: OLIGO

Xgapop 60.0	Xgapext 60.0
Ygapop 60.0	Ygapext 60.0
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 4708233 segs, 24227607955 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1429811

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-Processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n -DBV=rlp  
-G=/cgn2.1/USPTO.spool\_p/US10031158/runat\_14102005\_135842\_19603/aps\_query\_faeta.1.199  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n -OLIG\_SZ=30 -RGE -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bases -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=400 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HBAFSCORE=500 -MINLEN=0 -MAXLEN=30  
-USER=US10031158 @CCN 1.1 5600 @runat\_14102005\_135842\_19603 -NCPU=6 -ICPU=3  
-NOV\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DBV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEX=7

Database :

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sta:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
C 1	8	13.8	27	6 AX074409 Sequence
C 2	8	13.8	27	6 AX074410 Sequence
C 3	7	12.1	23	6 BD263082 Vector. 7
C 4	7	12.1	23	6 AX025028 Sequence

5	7	12.1	25	6	AX771994	Sequence
C 6	6	10.3	20	6	CO807420	Sequence
C 7	6	10.3	20	6	AR299570	Sequence
C 8	6	10.3	20	6	AX599128	Sequence
C 9	6	10.3	20	6	AX613466	Sequence
C 10	6	10.3	20	6	AX796046	Sequence
C 11	6	10.3	20	6	AX825279	Sequence
C 12	6	10.3	20	6	AX826219	Sequence
C 13	6	10.3	22	6	BD132408	Sequence
C 14	6	10.3	23	6	AX395174	Sequence
C 15	6	10.3	23	6	AX771993	Sequence
C 16	6	10.3	25	6	AX674984	Sequence
C 17	6	10.3	26	6	AX717020	Sequence
C 18	6	10.3	26	6	AX717041	Sequence
C 19	6	10.3	30	6	AR000072	Sequence
C 20	6	10.3	30	6	AR064895	Sequence
C 21	6	10.3	30	6	AX793255	Sequence
C 22	6	10.3	30	9	HUMPLTP26	Sequence
C 23	5	8.6	15	6	E49840	Novel phosph
C 24	5	8.6	15	6	BD083890	Novel phosph
C 25	5	8.6	15	6	BD096208	Novel phosph
C 26	5	8.6	16	6	AR002252	Sequence
C 27	5	8.6	16	6	AR045202	Sequence
C 28	5	8.6	16	6	AR051233	Sequence
C 29	5	8.6	16	6	AR066236	Sequence
C 30	5	8.6	16	6	AR101749	Sequence
C 31	5	8.6	16	6	AR127780	Sequence
C 32	5	8.6	16	6	AR127795	Sequence
C 33	5	8.6	16	6	AR166238	Sequence
C 34	5	8.6	16	6	BD263078	Novel
C 35	5	8.6	16	6	BD273667	Novel
C 36	5	8.6	16	6	116027	Sequence
C 37	5	8.6	16	6	128362	Sequence
C 38	5	8.6	16	6	128568	Sequence
C 39	5	8.6	16	6	128569	Sequence
C 40	5	8.6	16	6	158730	Sequence
C 41	5	8.6	16	6	158731	Sequence
C 42	5	8.6	16	6	AX025024	Sequence
C 43	5	8.6	16	6	AX769906	Sequence
C 44	5	8.6	17	6	A88667	Sequence
C 45	5	8.6	17	6	A90634	Sequence

#### ALIGNMENTS

RESULT 1  
AX074409/c  
LOCUS AX074409 27 bp DNA linear PAT 06-FEB-2001  
DEFINITION Sequence 7 from Patent WO0104309.  
ACCESSION AX074409  
VERSION AX074409.1 GI:12710547  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
other sequences; artificial sequences.  
REFERENCE  
1 Pastan, I., Esgand, M., Lee, B., Vasmatazis, G. and Wolfgang, C.  
T-cell receptor \_g(g) alternate reading frame protein, (tarp) and  
uses thereof  
JOURNAL UNITED STATES GOVERNMENT (US)  
PATENT: WO 0104309-A 7 18-JAN-2001;  
TITLE Location/Qualifiers  
FEATURES  
1..27  
source  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="PCR primer"

ORIGIN  
Alignment Scores:  
Pred. No.: 3.79 Length: 27  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 13.79% Indels: 0  
DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x AX074409 (1-27)

QY 13 LeuGlnLeuLysGlnSer 20  
26 CTTCAATTCCTGAACAAAGCTCC 3

RESULT 2  
AX074410/c AX074410 27 bp DNA linear PAT 06-FEB-2001  
DEFINITION Sequence 8 from Patent WO0104309.  
ACCESSION AX074410  
VERSION AX074410.1 GI:12710548  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Pastan I., Basand M., Lee B., Vasmatazis G. and Wolfgang C.  
TITLE T-cell receptor  $\gamma$ (g) alternate reading frame protein, (tarp) and uses thereof  
JOURNAL Patent: WO 0104309-A 8 18-JAN-2001;  
UNITED STATES GOVERNMENT (US)  
FEATURES  
source 1. 27  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="PCR primer"

ORIGIN  
Alignment Scores:  
Pred. No.: 3.79 Length: 27  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 13.79% Indels: 0  
DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x AX074410 (1-27)

QY 23 LeuGlnHleThrPheValPheLeu 30  
26 CTGGAACATACCTTTGCTTCTTG 3

RESULT 3  
BD263082/c BD263082 23 bp DNA linear PAT 17-JUL-2003  
LOCUS Vector.  
DEFINITION BD263082  
ACCESSION BD263082  
VERSION BD263082.1 GI:33072850  
KEYWORDS JP 2002530115-A/14.  
SOURCE Human immunodeficiency virus 1 (HIV-1)  
ORGANISM Human immunodeficiency virus 1  
Viruses; Retrovirdae; Retroviridae; Lentivirus; Primate  
lentivirus group.  
1 (bases 1 to 23)  
Vector  
Vector  
Patent: JP 2002530115-A 14 17-SEP-2002;  
OXFORD BIOMEDICA LTD  
OS Human immunodeficiency virus type 1  
PN JP 2002530115-A/14  
PD 17-SEP-2002  
PF 19-NOV-1999 JP 2000584089  
PI 20-NOV-1998 GB 9825524.3  
PI KYRIACOS MITROPHANOUS, MARK UDEN, JONATHAN ROHLT, SUSAN MARY PI  
KINGSMAN,  
PI ALAN KINGSMAN  
PC C12N15/09, A61K35/76, A61K48/00, A61P1/04, A61P9/00, A61P11/06, PC

A61P17/00,  
PC A61P25/00, A61P25/28, A61P27/02, A61P29/00, A61P31/12, A61P35/00,  
PC A61P37/00,  
PC C12N5/10, C12N7/00, C12N15/00, C12N5/00  
CC Vector  
FH Key  
FT source  
FT 1. 23  
FT Location/Qualifiers  
1.  
FT location/Qualifiers  
1. 23  
/organism="Human immunodeficiency virus 1"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:11676"

ORIGIN  
Alignment Scores:  
Pred. No.: 38 Length: 23  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.07% Indels: 0  
DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x BD263082 (1-23)

QY 5 ProProSerProLeuPhe 11  
22 CCCCATCCCCCTTTCTTT 2

RESULT 4  
AX025028/c AX025028 23 bp DNA linear PAT 15-SEP-2000  
LOCUS Sequence 14 from Patent WO0031280.  
DEFINITION AX025028  
ACCESSION AX025028  
VERSION AX025028.1 GI:10184948  
KEYWORDS  
SOURCE Human immunodeficiency virus 1 (HIV-1)  
ORGANISM Human immunodeficiency virus 1  
Viruses; Retrovirdae; Retroviridae; Lentivirus; Primate  
lentivirus group.

REFERENCE 1  
AUTHORS Kingsman, S.M., Mitrophanous, K., Uden, M., Rohlt, J. and Kingman, A.J.  
TITLE Vector  
JOURNAL Patent: WO 0031280-A 14 02-JUN-2000;  
KINGSMAN SUSAN MARY (GB); MITROPHANOUS KYRIACOS (GB); UDEN MARK  
(GB); ROHLT, JONATHAN (GB); KINGSMAN ALAN JOHN (GB); OXFORD  
BIOMEDICA LTD (GB)  
FEATURES  
source 1. 23  
/organism="Human immunodeficiency virus 1"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:11676"

ORIGIN  
Alignment Scores:  
Pred. No.: 38 Length: 23  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.07% Indels: 0  
DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x AX025028 (1-23)

QY 5 ProProSerProLeuPhe 11  
22 CCCCATCCCCCTTTCTTT 2

RESULT 5  
AX771994/c AX771994 25 bp DNA linear PAT 02-JUL-2003  
LOCUS Sequence 9 from Patent WO03006065.  
DEFINITION

ACCESSION AX771994  
 VERSION AX771994.1 GI:32438642  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 other sequences; artificial sequences.  
 REFERENCE 1  
 AUTHORS Braun, K., Waldeck, W., Pipkorn, R., Braun, I., and Debus, J.  
 TITLE PNA conjugate for the treatment of diseases associated with HIV  
 JOURNAL Patent: WO 0306065-A 9 23-JAN-2003;  
 Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts (DE)  
 FEATURES  
 source Location/Qualifiers  
 1..25  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="PNA Ia"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 41.1 Length: 25  
 Score: 7.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 12.07% Indels: 0  
 Gaps: 0  
 DB: 6  
 US-10-031-158-14 (1-58) x AX771994 (1-25)  
 QY 7 SerProLeuphPhePheLeu 13  
 DB 4 TCCCCCTTTCTTTTCTT 24  
 RESULT 6  
 LOCUS CQ807420 20 bp DNA linear PAT 10-MAY-2004  
 DEFINITION Sequence 870 from Patent WO2004035803.  
 ACCESSION CQ807420  
 VERSION CQ807420.1 GI:47112814  
 KEYWORDS  
 ORGANISM synthetic construct  
 SOURCE synthetic construct  
 other sequences; artificial sequences.  
 REFERENCE 1  
 AUTHORS Poekens, J., Harbeck, N., Koenig, T., Maier, S., Martens, J., Model, F., Nimrich, I., Rujan, T., Schmitt, A., Schmitt, M., Look, M. P. and Marx, A.  
 TITLE Method and nucleic acids for the improved treatment of breast cell proliferative disorders  
 JOURNAL Patent: WO 2004035803-A 870 29-APR-2004;  
 Epigenomics AG (DE)  
 FEATURES  
 source Location/Qualifiers  
 1..20  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="Detection primer for SFN"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 389 Length: 20  
 Score: 6.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.34% Indels: 0  
 Gaps: 0  
 DB: 6  
 US-10-031-158-14 (1-58) x CQ807420 (1-20)  
 QY 5 ProProSerProLeuphe 10  
 DB 18 CCTCCCTCTCTCTCTC 1

RESULT 7  
 LOCUS AR299570 20 bp DNA linear PAT 12-JUN-2003  
 DEFINITION Sequence 11305 from patent US 6537751.  
 ACCESSION AR299570  
 VERSION AR299570.1 GI:31686854  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 20)  
 AUTHORS Cohen, D., Chumakov, I., and Blumenfeld, M.  
 TITLE Biallelic markers for use in constructing a high density disequilibrium map of the human genome  
 JOURNAL Patent: US 6537751-A 11305 25-MAR-2003;  
 Location/Qualifiers  
 1..20  
 /organism="unknown"  
 /mol\_type="genomic DNA"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 389 Length: 20  
 Score: 6.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.34% Indels: 0  
 Gaps: 0  
 DB: 6  
 US-10-031-158-14 (1-58) x AR299570 (1-20)  
 QY 8 ProLeuphPhePheLeu 13  
 DB 3 CCCCTATTTTCTCTG 20  
 RESULT 8  
 LOCUS AX599128 20 bp DNA linear PAT 14-FEB-2003  
 DEFINITION Sequence 468 from Patent WO02077272.  
 ACCESSION AX599128  
 VERSION AX599128.1 GI:28399270  
 KEYWORDS  
 ORGANISM synthetic construct  
 SOURCE synthetic construct  
 other sequences; artificial sequences.  
 REFERENCE 1  
 AUTHORS Berlin, K., Braun, A., Discler, J., Guetig, D., Howe, A., Mueller, J., Olek, A., Piepenbrock, C., Adorian, P., Grabs, G., Lesche, R., Liu, B., Lewin, A., Lipcher, E., Maier, S., Model, F., Mueller, V., Otto, T., Pelet, C. and Ziebarth, H.  
 TITLE Methods and nucleic acids for the analysis of hematopoietic cell proliferative disorders  
 JOURNAL Patent: WO 02077272-A 468 03-OCT-2002;  
 Epigenomics AG (DE)  
 FEATURES  
 source Location/Qualifiers  
 1..20  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="Detection primer for SFN"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 369 Length: 20  
 Score: 6.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.34% Indels: 0  
 Gaps: 0  
 DB: 6  
 US-10-031-158-14 (1-58) x AX599128 (1-20)

Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	source	ORIGIN	
5	ProFroSerProLeuphe 10	Sequence 4491 from Patent WO02072882.	AX613466	AX613466.1	GI:28408895	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	1	Cullen, P. and Seedorf, U.	Coronary chip	Patent: WO 02072882-A 4491 19-SEP-2002;	OGHAM GmbH (DE)	Location/Qualifiers 1..20 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"	Alignment Scores: Pred. No.: 389 Score: 6.00 Percent Similarity: 100.00% Best Local Similarity: 100.00% Query Match: 10.34% DB: 6	Length: 20 Matches: 6 Conservative: 0 Mismatches: 0 Indels: 0 Gaps: 0
18	CCTCCCTCTCTCTCTTC 1	Sequence 389 from Patent WO03052135.	AX796046	AX796046.1	GI:37516712	synthetic construct synthetic construct other sequences; artificial sequences.	Burger, M., Field, J.K., Genc, B., Illoglou, T., Lipscher, E., Water, S. and Nimrich, I.	Method and nucleic acids for the analysis of a lung cell proliferative disorder	Patent: WO 03052135-A 389 26-JUN-2003;	Epigenomics AG (DE)	Location/Qualifiers 1..20 /organism="synthetic construct" /mol_type="unassigned DNA" /db_xref="taxon:32630" /note="Detection primer for SFN"	Alignment Scores: Pred. No.: 389 Score: 6.00 Percent Similarity: 100.00% Best Local Similarity: 10.34% DB: 6	Length: 20 Matches: 6 Conservative: 0 Mismatches: 0 Indels: 0 Gaps: 0		

```

US-10-031-158-14 (1-58) x AX796046 (1-20)

QY      5 ProProSerProLeuphe 10
DB      18 CCTCCTCTCTCTCTCTTC 1

RESULT 11
LOCUS   AX822579/c          20 bp      DNA      linear      PAT 11-DEC-2003
DEFINITION Sequence 471 from Patent EP1340818.
ACCESSION AX822579
VERSION   AX822579.1 GI:39749215
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Adorjan,P., Burger,M., Mater,S., Nimrich,I., Becker,E., Leasche,R.,
TITLE    Method and nucleic acids for the analysis of a colon cell
JOURNAL  Patent: EP 1340818-A 471 03-SEP-2003;
FEATURES
source   1..20
         /organism="synthetic construct"
         /mol_type="unassigned DNA"
         /db_xref="taxon:32630"
         /note="Detection primer for SFN"

ORIGIN
Alignment Scores:
Pred. No.:      389      Length:      20
Score:          6.00      Matches:      6
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches:  0
Query Match:    10.34%      Indels:      0
DB:             6      Gaps:      0

US-10-031-158-14 (1-58) x AX822579 (1-20)

QY      5 ProProSerProLeuphe 10
DB      18 CCTCCTCTCTCTCTCTTC 1

RESULT 12
LOCUS   AX826219/c          20 bp      DNA      linear      PAT 11-DEC-2003
DEFINITION Sequence 471 from Patent WO03072821.
ACCESSION AX826219
VERSION   AX826219.1 GI:39751733
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Adorjan,P., Burger,M., Mater,S., Nimrich,I., Becker,E., Leasche,R.,
TITLE    Method and nucleic acids for the analysis of a colon cell
JOURNAL  Patent: WO 03072821-A 471 04-SEP-2003;
FEATURES
source   1..20
         /organism="synthetic construct"
         /mol_type="unassigned DNA"
         /db_xref="taxon:32630"
         /note="Detection primer for SFN"

ORIGIN
Alignment Scores:
Pred. No.:      389      Length:      20
Score:          6.00      Matches:      6
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches:  0
Query Match:    10.34%      Indels:      0
DB:             6      Gaps:      0

```

Score: 6.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 10.34%  
 DB: 6 Matches: 6  
 Conservative: 0  
 Mismatches: 0  
 Indels: 0  
 Gaps: 0

US-10-031-158-14 (1-58) x AX826219 (1-20)

QY 5 ProProSerProleuphe 10  
 DB 18 CTTCTCTCTCTCTCTTC 1

RESULT 13  
 LOCUS BD132408/c 22 bp DNA linear PAT 18-SEP-2002  
 DEFINITION A basal cell carcinoma tumor suppressor gene.  
 ACCESSION BD132408  
 VERSION BD132408.1 GI:23227353  
 KEYWORDS JP 2002504805-A/20.  
 SOURCE synthetic construct  
 ORGANISM other sequences; artificial sequences.  
 REFERENCE 1 (bases 1 to 22)  
 AUTHORS Dean,M.F., Hahn,H., Wicking,C., Christiansen,J., Zaphiropoulos,P.G., Gallani,M.R., Shanley,S., Chidambaram,A., Vorechovsky,I., Holmberg,E., Unden,A.B., Gillies,S., Negus,K., Smyth,I., Pressman,C., Leffell,D.J., Gerrard,B., Goldstein,A., Wainwright,B., Tottgard,R., Trench,G.C. and Bale,A.E.  
 Patent: JP 2002504805-A 20 12-FEB-2002;  
 THE GOVERNMENT OF THE UNITED STATES OF AMERICA REPRESENTED BY THE SECRETARY DEPARTMENT OF HEALTH AND HUMAN SERVICES  
 PN JP 2002504805-A/20  
 PD 12-FEB-2002  
 PR 16-MAY-1997 JP 1997541164 PO 0011 PR  
 PR 17-MAY-1996 US 60/017906,21-MAY-1996 AU  
 07-JUN-1996 AU PO 0363,14-JUN-1996 US 60/019765 PI  
 MICHAEL, FREDERICK DEAN, HEIDI HAHN, CAROL WICKING, JEFFREY PI  
 CHRISTIANSEN,  
 PI PETER G ZAPHIROPOULOS, MAE R GALLANI, SUSAN SHANLEY, ABIRAMI PI  
 CHIDAMBARAM,  
 PI IGOR VORECHOVSKY, ERIKA HOLMBERG, ANNE BIRGITTE UNDEN, SUSAN PI  
 GILLIES,  
 PI KYLIE NEGUS, IAN SMYTH, CAROL PRESSMAN, DAVID J LEFFELL, BERNARD  
 GERARD,  
 PI ALISA GOLDSTEIN, BRANDON WAINWRIGHT, RONE TOTTGARD, GEORGIA PI  
 CHENEVIX TRENCHE,  
 PI ALLEN E BALE  
 PC C12N15/12, C07K14/47, C12N5/10, C12Q1/68, G01N33/50, A61K48/00, PC  
 A61K39/395,  
 PC A61K38/17  
 CC Strandedness: Single;  
 CC Topology: Linear;  
 CC /note='PTCF21 primer'  
 FH Key Location/Qualifiers

FEATURES  
 source 1..22  
 Location/Qualifiers  
 1..22  
 /organism="synthetic construct"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"

ORIGIN

Alignment Scores:  
 Pred. No.: 425 Length: 22  
 Score: 6.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.34% Indels: 0  
 DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x BD132408 (1-22)

QY 28 ValPheLeuArgAspHe 33

DB 21 GTGTCCTGGAATTTT 4

RESULT 14  
 LOCUS AX395174 23 bp DNA linear PAT 18-MAY-2002  
 DEFINITION Sequence 6 from Patent WO0218579.  
 ACCESSION AX395174  
 VERSION AX395174.1 GI:21066224  
 KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE 1  
 AUTHORS Glucksmann,M.A.  
 TITLE 57242, a novel human g protein-coupled receptor family member and  
 JOURNAL uses thereof  
 Patent: WO 0218579-A 6 07-MAR-2002;  
 Millennium Pharmaceuticals, Inc. (US)  
 Location/Qualifiers

FEATURES  
 source 1..23  
 Location/Qualifiers  
 1..23  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="murine 57242 probe sequence"

ORIGIN

Alignment Scores:  
 Pred. No.: 443 Length: 23  
 Score: 6.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.34% Indels: 0  
 DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x AX395174 (1-23)

QY 45 ArgArgAlaThrArgPhe 50  
 DB 4 AGGAGGCCACCCGGTTC 21

RESULT 15  
 LOCUS AX771993 23 bp DNA linear PAT 02-JUL-2003  
 DEFINITION Sequence 8 from Patent WO03006065.  
 ACCESSION AX771993  
 VERSION AX771993.1 GI:32438641  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE 1  
 AUTHORS Braun,K., Waldeck,W., Pipkorn,R., Braun,I. and Debus,J.  
 TITLE BNA conjugate for the treatment of diseases associated with HIV  
 JOURNAL Patent: WO 03006065-A 8 23-JAN-2003;  
 Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts  
 (DB)  
 Location/Qualifiers  
 1..23  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="PNA 1"

ORIGIN

Alignment Scores:  
 Pred. No.: 443 Length: 23  
 Score: 6.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.34% Indels: 0  
 DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x AX771993 (1-23)

Qy	7	SerProLeuPhePhePhe	12
Db	1	TCCCCCCTTTCTTTT	18

Search completed: October 16, 2005, 10:10:35  
Job time : 1777 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 16, 2005, 05:59:35 ; Search time 113 Seconds  
(without alignments)  
213.891 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 58

Sequence: 1 MQMPPSPPLFFFLQLKQSS.....RYGKRRATRFWDPRGTP 58

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1859788 seqs, 416717961 residues

Word size : 0

Total number of hits satisfying chosen parameters: 230707

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10F\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10G\_PUBCOMB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	8.6	5	US-09-895-443-30	Sequence 30, Appl
2	5	8.6	7	US-09-833-067-26	Sequence 26, Appl
3	5	8.6	8	US-10-137-867-292	Sequence 292, App
4	5	8.6	9	US-09-765-527-221	Sequence 221, App
5	5	8.6	9	US-09-792-480-18	Sequence 18, Appl
6	5	8.6	9	US-09-870-216C-3	Sequence 3, Appli
7	5	8.6	9	US-10-017-337-3	Sequence 2, Appli
8	5	8.6	9	US-10-144-188-2	Sequence 2, Appli
9	5	8.6	10	US-09-839-329-4	Sequence 4, Appli
10	5	8.6	10	US-09-572-404B-238	Sequence 238, App
11	5	8.6	10	US-09-572-404B-1444	Sequence 1444, Ap

12	5	8.6	10	US-09-572-404B-3212	Sequence 3212, Ap
13	5	8.6	10	US-10-007-761-9	Sequence 9, Appli
14	5	8.6	10	US-10-083-960-18	Sequence 18, Appl
15	5	8.6	10	US-10-209-421-18	Sequence 26, Appl
16	5	8.6	10	US-10-168-789A-26	Sequence 26, Appl
17	5	8.6	10	US-10-168-789A-27	Sequence 7, Appli
18	5	8.6	10	US-10-421-548-7	Sequence 65, Appl
19	5	8.6	10	US-10-421-503-65	Sequence 9, Appli
20	5	8.6	10	US-10-843-711-9	Sequence 47, Appl
21	5	8.6	10	US-10-936-237-47	Sequence 65, Appl
22	5	8.6	10	US-10-513-003-65	Sequence 65, Appl
23	4	6.9	4	US-09-793-451-717	Sequence 717, App
24	4	6.9	4	US-10-062-109A-701	Sequence 701, App
25	4	6.9	4	US-10-005-480A-701	Sequence 701, App
26	4	6.9	4	US-10-283-722-717	Sequence 717, App
27	4	6.9	4	US-10-291-241-41	Sequence 41, Appl
28	4	6.9	4	US-10-283-903-717	Sequence 717, App
29	4	6.9	4	US-10-133-234A-9	Sequence 9, Appli
30	4	6.9	4	US-10-418-032-253	Sequence 253, App
31	4	6.9	4	US-10-892-402-46	Sequence 36, Appl
32	4	6.9	5	US-09-895-443-24	Sequence 24, Appl
33	4	6.9	5	US-09-895-443-25	Sequence 25, Appl
34	4	6.9	5	US-09-895-443-26	Sequence 26, Appl
35	4	6.9	5	US-09-781-133-20	Sequence 20, Appl
36	4	6.9	5	US-09-781-133-21	Sequence 21, Appl
37	4	6.9	5	US-09-781-133-22	Sequence 22, Appl
38	4	6.9	5	US-10-235-483-29	Sequence 29, Appl
39	4	6.9	5	US-10-299-540-2	Sequence 2, Appli
40	4	6.9	5	US-10-299-184-2	Sequence 2, Appli
41	4	6.9	5	US-10-703-206-3	Sequence 3, Appli
42	4	6.9	5	US-10-808-187-1075	Sequence 1075, Ap
43	4	6.9	5	US-10-497-628-5	Sequence 5, Appli
44	4	6.9	5	US-10-776-224-04	Sequence 204, App
45	4	6.9	5	US-10-807-807-1075	Sequence 1075, Ap

#### ALIGNMENTS

RESULT 1  
US-09-895-443-30  
; Sequence 30, Application US/09895443  
; Patent No. US20020103134A1  
; GENERAL INFORMATION:  
; APPLICANT: Rindels, M. et al.  
; TITLE OF INVENTION: Aggregators of -Amyloid Peptide  
; CORRESPONDENCE ADDRESSES: 34  
; ADDRESSER: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1784  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Releasee #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/895,443  
; FILING DATE: 29-Jun-2001  
; CLASSIFICATION: <Unknown>  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 09/356,931  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: 08/897,342  
; FILING DATE: 21-JUL-1997  
; APPLICATION NUMBER: USSN 08/703,675  
; FILING DATE: 27-AUG-1996  
; APPLICATION NUMBER: USSN 08/616,081  
; FILING DATE: 14-MAR-1996

APPLICATION NUMBER: USSN 08/548,998  
FILING DATE: 27-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: DeConti, Jr., Giulio A.  
REGISTRATION NUMBER: 31,503  
REFERENCE/DOCKET NUMBER: PRI-016CP4CN  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified site  
LOCATION: 1-5  
OTHER INFORMATION: /note= D amino acid  
SEQUENCE DESCRIPTION: SEQ ID NO: 30:  
US-09-895-443-30

Query Match 8.6%; Score 5; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 LEFFL 13  
Db 1 LEFFL 5

RESULT 2  
US-09-833-067-26  
Sequence 26, Application US/09833067  
Patent No. US20020054888A1  
GENERAL INFORMATION:  
APPLICANT: O'HANLEY, KENNER  
APPLICANT: DENICH, KENNETH  
TITLE OF INVENTION: DISSOCIATED PILI, THEIR PRODUCTION AND USE  
FILE REFERENCE: 050939/0102  
CURRENT APPLICATION NUMBER: US/09/833,067  
CURRENT FILING DATE: 2001-04-12  
PRIOR APPLICATION NUMBER: 60/196,493  
PRIOR FILING DATE: 2000-04-12  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 26  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Unknown Sequence  
FEATURE:  
OTHER INFORMATION: Description of Unknown Sequence: highly hydrophobic region  
US-09-833-067-26

Query Match 8.6%; Score 5; DB 9; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 PLPFF 12  
Db 3 PLPFF 7

RESULT 3  
US-10-137-867-292  
Sequence 292, Application US/10137867  
Publication No. US20030207349A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C146  
CURRENT APPLICATION NUMBER: US/10/137,867  
CURRENT FILING DATE: 2002-05-03  
PRIOR APPLICATION removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 292  
LENGTH: 428  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-137-867-292

Query Match 8.6%; Score 5; DB 15; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 LEFFL 13  
Db 1 LEFFL 5

RESULT 4  
US-09-765-527-221  
Sequence 221, Application US/09765527  
Patent No. US2002000638A1  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
TITLE OF INVENTION: Methode for Recombinant Microbial Production of  
CORRESPONDENCE ADDRESS:  
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/765,527  
FILING DATE: 18-Jan-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/621,803  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Borun, Michael P.  
REGISTRATION NUMBER: 25,447  
REFERENCE/DOCKET NUMBER: 27129/33199  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 221:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids



TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: "IMP.382:  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: C-Terminus  
OTHER INFORMATION: /label= Amidation  
/note="The C-Terminus is Amidated."  
SEQUENCE DESCRIPTION: SEQ ID NO: 221:  
US-09-765-527-221

Query Match 8.6%; Score 5; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred.No.1.7e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LQLLK 17  
DB 4 LQLLK 8

RESULT 5  
US-09-792-480-18  
Sequence 18, Application US/09792480  
Patent No. US20020127198A1  
GENERAL INFORMATION:  
APPLICANT: Rothbard, Jonathan B.  
APPLICANT: Wender, Paul A.  
APPLICANT: McGrane, P. Leo  
APPLICANT: Sista, Lalitha V.S.  
APPLICANT: Kirschberg, Thorsten A.  
APPLICANT: Cellgate, Inc.  
TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery  
TITLE OF INVENTION: Actos and into Epithelial tissues  
FILE REFERENCE: 019801-000230US  
CURRENT FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: US 09/648,400  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/150,510  
PRIOR FILING DATE: 1999-08-24  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 18  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:A-54  
OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic  
OTHER INFORMATION: region tat-49-57  
NAME/KEY: MOD\_RES  
LOCATION: (1)  
OTHER INFORMATION: Xaa = fluorescein linked to amino group of  
OTHER INFORMATION: aminohexanoic acid (Fl-ahx) attached to the  
OTHER INFORMATION: N-terminal amino group of Arg  
US-09-792-480-18

Query Match 8.6%; Score 5; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred.No.1.7e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KGRRA 47  
DB 2 KGRRA 6

RESULT 6  
US-09-870-216C-3  
Sequence 3, Application US/09870216C  
Publication No. US20040138135A1

GENERAL INFORMATION:  
APPLICANT: Charles A. Nicolette  
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER  
FILE REFERENCE: 68126881210100  
CURRENT APPLICATION NUMBER: US/09/870,216C  
CURRENT FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: 60/209,391  
PRIOR FILING DATE: 2000-05-31  
PRIOR APPLICATION NUMBER: 60/226,256  
PRIOR FILING DATE: 2000-08-17  
PRIOR APPLICATION NUMBER: 60/257,008  
PRIOR FILING DATE: 2000-12-20  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-870-216C-3

Query Match 8.6%; Score 5; DB 11; Length 9;  
Best Local Similarity 100.0%; Pred.No.1.7e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FLQLL 16  
DB 1 FLQLL 5

RESULT 7  
US-10-017-327-3  
Sequence 3, Application US/10017327  
Publication No. US20020155471A1  
GENERAL INFORMATION:  
APPLICANT: Charles A. Nicolette  
TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS AND  
TITLE OF INVENTION: METHODS FOR USING SAME  
FILE REFERENCE: GZ 2101.20  
CURRENT APPLICATION NUMBER: US/10/017,327  
CURRENT FILING DATE: 2001-12-06  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-017-327-3

Query Match 8.6%; Score 5; DB 13; Length 9;  
Best Local Similarity 100.0%; Pred.No.1.7e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FLQLL 16  
DB 1 FLQLL 5

RESULT 8  
US-10-144-188-2  
Sequence 2, Application US/10144188  
Publication No. US20030170212A1  
GENERAL INFORMATION:  
APPLICANT: Cai, Zeling  
APPLICANT: Jackson, Michael R.  
APPLICANT: Peterson, Per A.  
APPLICANT: Shi, Weixing  
APPLICANT: Kong, Yan  
APPLICANT: Degraw, Juli  
TITLE OF INVENTION: Ex-Vivo Priming For Generating Cytotoxic T Lymphocytes Specific  
TITLE OF INVENTION: For No. US20030170212A1-Tumor Antigens To Treat Autoimmune And  
FILE REFERENCE: PRI0010 ORI-1627  
CURRENT APPLICATION NUMBER: US/10/144,188  
CURRENT FILING DATE: 2002-05-13

```

; PRIOR APPLICATION NUMBER: 60/291,300
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide antigen
US-10-144-188-2

Query Match      8.6%; Score 5; DB 14; Length 9;
Best Local Similarity 100.0%; Pred.No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 PPSPL 9
        |||||
DB      2 PPSPL 6

RESULT 9
US-09-839-329-4
; Sequence 4, Application US/09839329
; Publication No. US20020022027A1
; GENERAL INFORMATION:
; APPLICANT: Michael P. Sherman
; APPLICANT: Warner C. Greene
; APPLICANT: Carlos M.C. de No. 66640400hna
; APPLICANT: Ulrich Schubert
; APPLICANT: Peter Henklein
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DELIVERY OF
; FILE REFERENCE: A MOLECULE INTO A CELL
; CURRENT APPLICATION NUMBER: US/09/839,329
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/206,610
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/267,827
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human immunodeficiency virus
US-09-839-329-4

Query Match      8.6%; Score 5; DB 9; Length 10;
Best Local Similarity 100.0%; Pred.No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      42 GKRR 46
        |||||
DB      2 GKRR 6

RESULT 10
US-09-572-404B-238
; Sequence 238, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 238
; LENGTH: 10
```

```

; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in MTRR1B at 351-360 and may interact with Sequ
US-09-572-404B-238

Query Match      8.6%; Score 5; DB 10; Length 10;
Best Local Similarity 100.0%; Pred.No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 SRRL 24
        |||||
DB      1 SRRL 5

RESULT 11
US-09-572-404B-1444
; Sequence 1444, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 1444
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in NTRK at 681-690 and may interact with Sequenc
US-09-572-404B-1444

Query Match      8.6%; Score 5; DB 10; Length 10;
Best Local Similarity 100.0%; Pred.No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      53 PRRG 57
        |||||
DB      6 PRRG 10

RESULT 12
US-09-572-404B-3212
; Sequence 3212, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 3212
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in HOXA9 OR HOX1G at 369-378 and may interact w
US-09-572-404B-3212

Query Match      8.6%; Score 5; DB 10; Length 10;
Best Local Similarity 100.0%; Pred.No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      45 RRATR 49
        |||||
```

DB 4 RRATR 8

## RESULT 13

US-10-007-761-9  
Sequence 9, Application US/10007761  
Publication No. US20020150984A1  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: Peptides for Activation and Inhibition  
FILE REFERENCE: 58600-8208, US00  
CURRENT FILING DATE: 2001-11-09  
PRIOR APPLICATION NUMBER: US 60/262,060  
PRIOR FILING DATE: 2001-01-18  
NUMBER OF SEQ ID NOS: 72  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Tat-derived carrier peptide  
US-10-007-761-9

Query Match 8.6%; Score 5; DB 13; Length 10;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 GKKRR 46  
|||||  
DB 2 GKKRR 6

## RESULT 14

US-10-083-960-18  
Sequence 18, Application US/10083960  
Publication No. US20030022831A1  
GENERAL INFORMATION:  
APPLICANT: Rothbard, Jonathan B.  
APPLICANT: Wender, Paul A.  
APPLICANT: McGrane, P. Leo  
APPLICANT: Sista, Lalitha V.S.  
APPLICANT: Kirschberg, Thorsten A.  
TITLE OF INVENTION: Compositions and Methods for Enhancing  
FILE REFERENCE: 019801-000240US  
CURRENT FILING DATE: 2003-07-14  
PRIOR APPLICATION NUMBER: US 60/150,510  
PRIOR FILING DATE: 1999-08-24  
PRIOR APPLICATION NUMBER: US 09/648,400  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 09/792,480  
PRIOR FILING DATE: 2001-02-23  
NUMBER OF SEQ ID NOS: 86  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 18  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: A-54 alanine-substituted analog of Tat-49-57  
NAME/KEY: MOD\_RES  
LOCATION: (1)...(1)  
OTHER INFORMATION: Xaa = fluorescein conjugated aminohexanoic acid  
US-10-083-960-18

Query Match 8.6%; Score 5; DB 14; Length 10;

Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKRRA 47  
|||||  
DB 3 KKRRA 7

## RESULT 15

US-10-209-421-18  
Sequence 18, Application US/10209421  
Publication No. US20030083256A1  
GENERAL INFORMATION:  
APPLICANT: Rothbard, Jonathan B.  
APPLICANT: Wender, Paul A.  
APPLICANT: McGrane, P. Leo  
APPLICANT: Sista, Lalitha V.S.  
APPLICANT: Kirschberg, Thorsten A.  
TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery  
FILE REFERENCE: 019801-000211US  
CURRENT FILING DATE: 2002-07-30  
PRIOR APPLICATION NUMBER: US 60/150,510  
PRIOR FILING DATE: 1999-08-24  
PRIOR APPLICATION NUMBER: US 09/648,400  
PRIOR FILING DATE: 2000-08-24  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 18  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: A-54  
OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic  
FEATURE:  
NAME/KEY: MOD\_RES  
LOCATION: (1)  
OTHER INFORMATION: Xaa = fluorescein linked to amino group of  
US-10-209-421-18

Query Match 8.6%; Score 5; DB 14; Length 10;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKRRA 47  
|||||  
DB 3 KKRRA 7

Search completed: October 16, 2005, 06:11:00  
Job time: 114 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 16, 2005, 05:53:10 ; Search time 55 Seconds  
(without alignments)  
78.721 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 58  
Sequence: 1 MQMPPSPLEFFLQLLKQSS.....RYGKKRRATRPWDPRGTP 58

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 115750

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database : Issued Patents, AA:\*

1: /cgn2\_6/pdata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/pdata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/pdata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/pdata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/pdata/1/1aa/PCUS\_COMB.pep:\*  
6: /cgn2\_6/pdata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5	8.6	5	2	US-08-920-162A-30 Sequence 30, Appl
2	5	8.6	5	3	US-09-356-931-30 Sequence 30, Appl
3	5	8.6	5	4	US-09-895-443A-30 Sequence 30, Appl
4	5	8.6	7	4	US-09-627-851B-47 Sequence 47, Appl
5	5	8.6	9	2	US-08-621-803-221 Sequence 221, App
6	5	8.6	9	2	US-08-621-259A-213 Sequence 213, App
7	5	8.6	9	3	US-08-159-339A-343 Sequence 343, App
8	5	8.6	9	3	US-08-159-339A-405 Sequence 405, App
9	5	8.6	9	3	US-09-217-352-221 Sequence 221, App
10	5	8.6	9	4	US-09-677-664B-213 Sequence 213, App
11	5	8.6	9	4	US-09-792-480-18 Sequence 18, App
12	5	8.6	10	4	US-08-248-839C-135 Sequence 135, App
13	5	8.6	10	4	US-09-648-400A-18 Sequence 18, App
14	5	8.6	10	4	US-09-839-329-4 Sequence 4, Appl
15	5	8.6	10	4	US-10-209-421-18 Sequence 18, Appl
16	5	8.6	10	5	PCT-US96-01735-4 Sequence 4, Appl
17	4	6.9	4	1	US-07-917-034A-3 Sequence 3, Appl
18	4	6.9	4	3	US-09-085-305-29 Sequence 29, Appl
19	4	6.9	4	4	US-09-627-851B-21 Sequence 21, Appl
20	4	6.9	4	4	US-09-627-851B-29 Sequence 29, Appl
21	4	6.9	4	5	PCT-US93-01669-36 Sequence 36, Appl
22	4	6.9	5	1	US-07-690-284A-2 Sequence 2, Appl
23	4	6.9	5	1	US-07-690-284A-5 Sequence 5, Appl
24	4	6.9	5	1	US-08-704-170-91 Sequence 91, Appl
25	4	6.9	5	1	US-08-618-464-2 Sequence 2, Appl
26	4	6.9	5	2	US-08-656-906-28 Sequence 28, Appl
27	4	6.9	5	2	US-08-920-162A-24 Sequence 24, Appl

28	4	6.9	5	2	US-08-920-162A-25 Sequence 25, Appl
29	4	6.9	5	2	US-08-920-162A-26 Sequence 26, Appl
30	4	6.9	5	3	US-08-752-526-13 Sequence 13, Appl
31	4	6.9	5	3	US-09-107-615-2 Sequence 2, Appl
32	4	6.9	5	3	US-09-217-847-28 Sequence 28, Appl
33	4	6.9	5	3	US-09-356-931-24 Sequence 24, Appl
34	4	6.9	5	3	US-09-356-931-25 Sequence 25, Appl
35	4	6.9	5	3	US-09-356-931-25 Sequence 25, Appl
36	4	6.9	5	3	US-08-703-675C-24 Sequence 24, Appl
37	4	6.9	5	3	US-08-703-675C-25 Sequence 25, Appl
38	4	6.9	5	3	US-08-703-675C-26 Sequence 26, Appl
39	4	6.9	5	4	US-08-766-596A-29 Sequence 29, Appl
40	4	6.9	5	4	US-09-895-443A-24 Sequence 24, Appl
41	4	6.9	5	4	US-09-895-443A-25 Sequence 25, Appl
42	4	6.9	5	4	US-09-895-443A-26 Sequence 26, Appl
43	4	6.9	5	4	US-09-674-973A-204 Sequence 204, App
44	4	6.9	5	5	PCT-US94-02631-91 Sequence 91, Appl
45	4	6.9	5	5	PCT-US95-04896-2 Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-08-920-162A-30  
Sequence 30, Application US/08920162A  
Patent No. 5985242  
GENERAL INFORMATION:  
APPLICANT: FIndels, M. et al.  
TITLE OF INVENTION: Modulators of [SYMBOL] 98 \F "Symbol"-Amyloid Peptide  
NUMBER OF INVENTIONS: Aggregation Comprising D-Amino Acids  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: LAHIVE & COCKFIELD  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/920,162A  
FILING DATE: Herewith  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/548,998  
FILING DATE: 27-OCT-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/616,081  
FILING DATE: 14-MAR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/703,675  
FILING DATE: 27-AUG-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE: 21-JUL-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: KARA, Catherine J.  
REGISTRATION NUMBER: 41,106  
REFERENCE/DOCKET NUMBER: PPI-016CP4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

FEATURE:  
NAME/KEY: Modified site  
LOCATION: 1-5  
OTHER INFORMATION: /note= D amino acid  
US-08-920-162A-30

Query Match 8.6%; Score 5; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LFFFL 13  
|||||  
Db 1 LFFFL 5

RESULT 2  
US-09-356-931-30  
Sequence 30, Application US/09356931  
Patent No. 6277826

GENERAL INFORMATION:  
APPLICANT: Findels, M. et al.  
TITLE OF INVENTION: Modulators of -Amyloid Peptide  
TITLE OF INVENTION: Aggregation Compprising D-Amino Acids  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSES: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1784

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/356,931  
FILING DATE: 19-JUL-1999

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/920,162  
FILING DATE: 27-AUG-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/897,342  
FILING DATE: 21-JUL-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/703,675  
FILING DATE: 27-AUG-1996

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/616,081  
FILING DATE: 14-MAR-1996

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/548,998  
FILING DATE: 27-OCT-1995

ATTORNEY/AGENT INFORMATION:  
NAME: DeConti, Jr., Giulio A.  
REGISTRATION NUMBER: 31,503  
REFERENCE/DOCKET NUMBER: PPI-016CP4CN  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

FEATURE:  
NAME/KEY: Modified site  
LOCATION: 1-5  
OTHER INFORMATION: /note= D amino acid  
US-09-356-931-30

Query Match 8.6%; Score 5; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LFFFL 13  
|||||  
Db 1 LFFFL 5

RESULT 3  
US-09-895-443A-30  
Sequence 30, Application US/09895443A  
Patent No. 6689752

GENERAL INFORMATION:  
APPLICANT: FINDELS, MARK A.  
APPLICANT: GEFTER, MARCOLM L.  
APPLICANT: MUSCO, GARY  
APPLICANT: SIGNER, ETRIAN R.  
APPLICANT: WAKFIELD, JAMES  
APPLICANT: MOLINAUX, SUSAN  
APPLICANT: CHIN, JOSEPH  
APPLICANT: LEE, JUNG-JA  
APPLICANT: KELLEY, MICHAEL  
APPLICANT: KOMAR-PANICUCCI, SONJA  
APPLICANT: ARICO-MUENDEL, CHRISTOPHER C.  
APPLICANT: PHILLIPS, KATHRYN  
APPLICANT: HAYWARD, NEIL J.

TITLE OF INVENTION: MODULATORS OF BETA-AMYLOID PEPTIDE  
TITLE OF INVENTION: AGGREGATION COMPRISING D-AMINO ACIDS  
FILE REFERENCE: PPI-0016CP4CN2

CURRENT APPLICATION NUMBER: US/09/895,443A  
CURRENT FILING DATE: 2001-06-29

PRIOR APPLICATION NUMBER: 09/356931  
PRIOR FILING DATE: 1990-07-19

PRIOR APPLICATION NUMBER: 08/920162  
PRIOR FILING DATE: 1997-08-27

PRIOR APPLICATION NUMBER: 08/548998  
PRIOR FILING DATE: 1995-10-27

PRIOR APPLICATION NUMBER: 08/616081  
PRIOR FILING DATE: 1996-03-14

PRIOR APPLICATION NUMBER: 08/703675  
PRIOR FILING DATE: 1996-08-27

PRIOR APPLICATION NUMBER: 08/897342  
PRIOR FILING DATE: 1997-07-21

NUMBER OF SEQ ID NOS: 34  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 30

LENGTH: 5  
TYPE: PPT  
ORGANISM: Artificial Sequence

FEATURE:  
OTHER INFORMATION: peptide

FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(5)  
OTHER INFORMATION: Modified site, D amino acid  
US-09-895-443A-30

Query Match 8.6%; Score 5; DB 4; Length 5;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LFFFL 13  
|||||  
Db 1 LFFFL 5

RESULT 4  
US-09-627-851B-47  
Sequence 47, Application US/09627851B  
Patent No. 6545131

GENERAL INFORMATION:

APPLICANT: ISSACS, JOHN T.  
APPLICANT: DENNEADE, SAMUEL R.  
APPLICANT: CHRISTENSEN, S. BROGGER  
APPLICANT: LILJA, HANS  
TITLE OF INVENTION: TISSUE SPECIFIC PRODRUG  
FILE REFERENCE: 57111 (71699)  
CURRENT APPLICATION NUMBER: US/09/627,851B  
CURRENT FILING DATE: 2000-07-28  
PRIOR APPLICATION NUMBER: 09/588,822  
PRIOR FILING DATE: 2000-06-07  
PRIOR APPLICATION NUMBER: 09/081,707  
PRIOR FILING DATE: 1998-05-19  
PRIOR APPLICATION NUMBER: 60/047,070  
PRIOR FILING DATE: 1997-05-19  
PRIOR APPLICATION NUMBER: 60/080,046  
PRIOR FILING DATE: 1998-03-30  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO: 47  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-627-851B-47

Query Match 8.6%; Score 5; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 GKRR 46  
|||||  
Db 2 GKRR 6

RESULT 5  
US-08-621-803-221  
Sequence 221, Application US/08621803  
Patent No. 5851802  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
TITLE OF INVENTION: Methods for Recombinant Microbial Production of  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/621,803  
FILING DATE: 22-MAR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Borun, Michael F.  
REGISTRATION NUMBER: 25,447  
REFERENCE/DOCKET NUMBER: 27129/33199  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 221:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULAR TYPE: peptide  
FEATURE:  
NAME/KEY: misc\_feature

OTHER INFORMATION: "XMP.382:  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: C-Terminus  
OTHER INFORMATION: /label= Amidation  
OTHER INFORMATION: /note= "The C-Terminus is Amidated."  
US-08-621-803-221

Query Match 8.6%; Score 5; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LQLK 17  
|||||  
Db 4 LQLK 8

RESULT 6  
US-08-621-259A-213  
Sequence 213, Application US/08621259A  
Patent No. 5858974  
GENERAL INFORMATION:  
APPLICANT: Little II, Roger G  
APPLICANT: Lam, Edward  
APPLICANT: Fadum, Mitchell B.  
TITLE OF INVENTION: Anti-Fungal Peptides  
NUMBER OF SEQUENCES: 252  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/621,259A  
FILING DATE: 21-MAR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/504,841  
FILING DATE: 20-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 11021US02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX:  
INFORMATION FOR SEQ ID NO: 213:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULAR TYPE: peptide  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: "XMP.382:  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: C-Terminus  
OTHER INFORMATION: /label= Amidation  
OTHER INFORMATION: /note= "The C-Terminus is Amidated."  
US-08-621-259A-213

Query Match 8.6%; Score 5; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LQLK 17  
|||||  
DB 4 LQLK 8

RESULT 7  
US-08-159-339A-343

/ Sequence 343, Application US/08159339A  
/ Patent No. 6037135

/ GENERAL INFORMATION:

/ APPLICANT: Kubo, Ralph T.

/ APPLICANT: Sette, Alessandro

/ APPLICANT: Celis, Esteban

/ TITLE OF INVENTION: HLA Binding peptides and Their

/ NUMBER OF SEQUENCES: 1254

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: Townsend and Townsend and Crew LLP

/ STREET: Two Embarcadero Center, Eighth Floor

/ CITY: San Francisco

/ STATE: CA

/ COUNTRY: USA

/ ZIP: 94111-3834

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Diskette

/ COMPUTER: IBM Compatible

/ OPERATING SYSTEM: DOS

/ SOFTWARE: FASTSEQ for Windows Version 2.0

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/159,339A

/ FILING DATE: 29-NOV-1993

/ CLASSIFICATION: 424

/ PRIORITY APPLICATION DATA:

/ APPLICATION NUMBER: US 07/926,666

/ FILING DATE: 07-AUG-1992

/ APPLICATION NUMBER: US 08/027,746

/ FILING DATE: 05-MAR-1993

/ APPLICATION NUMBER: US 08/103,396

/ FILING DATE: 06-AUG-1993

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Weber, Ellen Lauver

/ REGISTRATION NUMBER: 32,762

/ REFERENCE/DOCKET NUMBER: 018623-005030US

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (415) 576-0200

/ TELEFAX: (415) 576-0300

/ TELEX:

/ INFORMATION FOR SEQ ID NO: 343:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 9 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: peptide

/ US-08-159-339A-343

/

/

/

/

/

/

/

/

/

/

/

/

/

/ APPLICANT: Sette, Alessandro

/ APPLICANT: Celis, Esteban

/ TITLE OF INVENTION: HLA Binding peptides and Their

/ NUMBER OF SEQUENCES: 1254

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: Townsend and Townsend and Crew LLP

/ STREET: Two Embarcadero Center, Eighth Floor

/ CITY: San Francisco

/ STATE: CA

/ COUNTRY: USA

/ ZIP: 94111-3834

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Diskette

/ COMPUTER: IBM Compatible

/ OPERATING SYSTEM: DOS

/ SOFTWARE: FASTSEQ for Windows Version 2.0

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/159,339A

/ FILING DATE: 29-NOV-1993

/ CLASSIFICATION: 424

/ PRIORITY APPLICATION DATA:

/ APPLICATION NUMBER: US 07/926,666

/ FILING DATE: 07-AUG-1992

/ APPLICATION NUMBER: US 08/027,746

/ FILING DATE: 05-MAR-1993

/ APPLICATION NUMBER: US 08/103,396

/ FILING DATE: 06-AUG-1993

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Weber, Ellen Lauver

/ REGISTRATION NUMBER: 32,762

/ REFERENCE/DOCKET NUMBER: 018623-005030US

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (415) 576-0200

/ TELEFAX: (415) 576-0300

/ TELEX:

/ INFORMATION FOR SEQ ID NO: 405:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 9 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: peptide

/ US-08-159-339A-405

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

QY 34 SIML 38  
|||||  
DB 3 SIML 7

RESULT 9  
US-09-217-352-221

/ Sequence 221, Application US/09217352

/ Patent No. 6274344

/ GENERAL INFORMATION:

/ APPLICANT: Better, Marc D.

/ APPLICANT: Methods for Recombinant Microbial Production of

/ TITLE OF INVENTION: Fusion Proteins and BPI-Derived Peptides

/ NUMBER OF SEQUENCES: 265

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum

/ STREET: 6300 Sears Tower, 233 South Wacker Drive

/ CITY: Chicago

/ STATE: Illinois

/ COUNTRY: United States of America

/ ZIP: 60606-6402

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/



OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/217,352  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/621,803  
FILING DATE: 22-MAR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Borun, Michael P.  
REGISTRATION NUMBER: 25,447  
REFERENCE/DOCKET NUMBER: 27129/33199  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 221:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: "XMP.382:"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: C-Terminus  
OTHER INFORMATION: /label= Amidation  
OTHER INFORMATION: /note= "The C-Terminus is Amidated."  
US-09-217-352-221

Query Match 8.6%; Score 5; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LQLK 17  
|||||  
Db 4 LQLK 8

RESULT 10  
US-09-677-664B-213  
Sequence 213, Application US/09677664B  
Patent No. 6664231  
GENERAL INFORMATION:  
APPLICANT: Little II, Roger G  
Liu, Edward  
Fadem, Mitchell B.  
TITLE OF INVENTION: Anti-Fungal Peptides  
NUMBER OF SEQUENCES: 257  
CORRESPONDENCE ADDRESS:  
ADDRESSER: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/677,664B  
FILING DATE: 07-Mar-2003  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/227,659  
FILING DATE: 08-Jan-1999  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 11021US06

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/775-8000  
TELEFAX: 312/775-8100  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 213:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: "XMP.382:"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: C-Terminus  
OTHER INFORMATION: /label= Amidation  
OTHER INFORMATION: /note= "The C-Terminus is Amidated."  
SEQUENCE DESCRIPTION: SEQ ID NO: 213:  
US-09-677-664B-213

Query Match 8.6%; Score 5; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LQLK 17  
|||||  
Db 4 LQLK 8

RESULT 11  
US-09-792-480-18  
Sequence 18, Application US/09792480  
Patent No. 6669951  
GENERAL INFORMATION:  
APPLICANT: Rothbard, Jonathan B.  
APPLICANT: Wender, Paul A.  
APPLICANT: McGrane, P. Leo  
APPLICANT: Siera, Lalitha V.S.  
APPLICANT: Kirschberg, Thoresten A.  
APPLICANT: Cellgate, Inc.  
TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery  
FILE REFERENCE: 019801-000230US  
CURRENT APPLICATION NUMBER: US/09/792,480  
CURRENT FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: US 09/648,400  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/150,510  
PRIOR FILING DATE: 1999-08-24  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 18  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:A-54  
OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic  
OTHER INFORMATION: region Tat-49-57  
NAME/KEY: MOD\_RES  
LOCATION: (1)  
OTHER INFORMATION: Xaa = fluorescein linked to amino group of  
OTHER INFORMATION: aminohexanoic acid (Fl-ahx) attached to the  
OTHER INFORMATION: N-terminal amino group of Arg  
US-09-792-480-18

Query Match 8.6%; Score 5; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 KRRR 47  
|||||

DB 2 KKRRA 6

## RESULT 12

US-08-248-839C-135  
Sequence 135, Application US/0824839C

GENERAL INFORMATION:  
PATENT NO. 5843702  
APPLICANT: McConnell, David  
APPLICANT: Devine, Kevin  
APPLICANT: O'Kane, Charles  
TITLE OF INVENTION: A Gene Expression System  
NUMBER OF SEQUENCES: 185  
CORRESPONDENCE ADDRESS:  
ADDRESS: No. 58437020 No. 5843702disk of No. 5843702th America, Inc.  
STREET: 405 Lexington Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/248, 839C  
FILING DATE: 25-MAY-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gregg, Valeta A.  
REGISTRATION NUMBER: 35,127  
REFERENCE/DOCKET NUMBER: 3614, 214-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 135:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
US-08-248-839C-135

Query Match 8.6%; Score 5; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 SIML 38

DB 3 SIML 7

## RESULT 13

US-09-648-400A-18  
Sequence 18, Application US/09648400A

PATENT NO. 6593292  
GENERAL INFORMATION:  
APPLICANT: Rothbard, Jonathan B.  
APPLICANT: Wender, Paul A.  
APPLICANT: McGrane, P. Leo  
APPLICANT: Sista, Lalitha V.S.  
APPLICANT: Kirschberg, Thorsten A.  
APPLICANT: Celigate, Inc.  
TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery  
TITLE OF INVENTION: Across and into Epithelial Tissues  
FILE REFERENCE: 019801-000210US  
CURRENT APPLICATION NUMBER: US/09/648,400A  
CURRENT FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/150,510  
PRIOR FILING DATE: 1999-08-24  
NUMBER OF SEQ ID NOS: 30

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 18  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: A-54  
OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic  
FEATURE:  
NAME/KEY: MOD RES  
LOCATION: (1)\_RES  
OTHER INFORMATION: Xaa = fluorescein linked to amino group of  
OTHER INFORMATION: aminohexanoic acid (Fl-ahx)  
US-09-648-400A-18

Query Match 8.6%; Score 5; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKRRA 47

DB 3 KKRRA 7

## RESULT 14

US-09-839-329-4  
Sequence 4, Application US/09839329

PATENT NO. 6664040  
GENERAL INFORMATION:  
APPLICANT: Michael P. Sherman  
APPLICANT: Warner C. Greene  
APPLICANT: Carlos M.C. de No. 6664040china  
APPLICANT: Ulrich Schubert  
APPLICANT: Peter Henklein  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DELIVERY OF  
TITLE OF INVENTION: A MOLECULE INTO A CELL  
FILE REFERENCE: GEC 30448, 91-US-02  
CURRENT APPLICATION NUMBER: US/09/839,329  
CURRENT FILING DATE: 2001-04-20  
PRIOR APPLICATION NUMBER: 60/206,610  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/267,827  
PRIOR FILING DATE: 2001-02-09  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Human immunodeficiency virus  
US-09-839-329-4

Query Match 8.6%; Score 5; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 GKRR 46

DB 2 GKRR 6

## RESULT 15

US-10-209-421-18  
Sequence 18, Application US/10209421

PATENT NO. 6759387  
GENERAL INFORMATION:  
APPLICANT: Rothbard, Jonathan B.  
APPLICANT: Wender, Paul A.  
APPLICANT: McGrane, P. Leo  
APPLICANT: Sista, Lalitha V.S.  
APPLICANT: Kirschberg, Thorsten A.

; APPLICANT: Cellgate, Inc.  
 ; TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery  
 ; FILE REFERENCE: 019801-000211US  
 ; CURRENT FILING DATE: 2002-07-30  
 ; PRIOR APPLICATION NUMBER: US 60/150,510  
 ; PRIOR FILING DATE: 1999-08-24  
 ; PRIOR APPLICATION NUMBER: US 09/648,400  
 ; PRIOR FILING DATE: 2000-08-24  
 ; NUMBER OF SEQ ID NOS: 51  
 ; SOFTWARE: Patent Ver. 2.1  
 ; SEQ ID NO: 18  
 ; LENGTH: 10  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence:A-54  
 ; OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic  
 ; OTHER INFORMATION: region Tat-49-57  
 ; NAME/KEY: MOD\_RES  
 ; LOCATION: (1)  
 ; OTHER INFORMATION: Xaa = fluorescein linked to amino group of  
 ; OTHER INFORMATION: aminohexanoic acid (Fl-ahx)  
 US-10-209-421-18

Query Match 8.6%; Score 5; DB 4; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 KGRRA 47  
 |||||  
 Db 3 KGRRA 7

Search completed: October 16, 2005, 06:01:38  
 Job time : 56 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: October 16, 2005, 05:45:39 ; Search time 118 Seconds

(without alignments)  
251.700 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 58  
Sequence: 1 MQMPPSPPLFFFLQLKQSS.....RYGKRRATRFMPDRGTP 58

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2548

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database : UniProt\_03:\*

1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4	6.9	4	1 FLRN ANTEL	P58707 antiopeura
2	4	6.9	7	P92210	P92210 agropyron c
3	4	6.9	7	P92214	P92214 amblyopyrum
4	4	6.9	7	P92218	P92218 auctrologyx
5	4	6.9	7	P92221	P92221 bromus iner
6	4	6.9	7	P92226	P92226 cithopsis
7	4	6.9	7	P92372	P92372 haynaldia v
8	4	6.9	7	P92381	P92381 hordium bra
9	4	6.9	7	P92385	P92385 hordium mar
10	4	6.9	7	P92387	P92387 hennardia p
11	4	6.9	7	P92390	P92390 heteranthel
12	4	6.9	7	P92393	P92393 hordium vul
13	4	6.9	7	P92403	P92403 lophopyrum
14	4	6.9	7	P92421	P92421 psathyrosta
15	4	6.9	7	P92425	P92425 pseudoroegn
16	4	6.9	7	P92427	P92427 peridictyon
17	4	6.9	7	P92430	P92430 aegilops ta
18	4	6.9	7	P92440	P92440 chinopyrum
19	4	6.9	7	P92442	P92442 taeniatheru
20	4	6.9	8	O15896	O15896 babesia bov
21	4	6.9	9	O88889	O88889 mus musculu
22	4	6.9	10	1 PARP LOCM1	P38553 locusta mig
23	4	6.9	10	1 LCMS_LEUMA	P21144 leucophaea
24	4	6.9	10	1 NEMS_SARBU	P61150 sarcophaga
25	4	6.9	10	2 O6A3T0	O6A3T0 archanglopt
26	4	6.9	10	2 O76ML8	O76ML8 eurypharynx
27	3	5.2	4	1 FLRN HRME	P42561 hirtudo medi
28	3	5.2	5	1 PARP_ARTTR	P41853 artiposethi
29	3	5.2	5	1 PAR2_PARMA	P81864 pardachilus
30	3	5.2	7	1 FARI_HELTI	P41871 helisma tr
31	3	5.2	7	1 FARI_MACRS	P83374 macrobrachi

32	3	5.2	7	1 FARI_PROCL	P38499 procamburus
33	3	5.2	7	1 FARI_PROCL	P38498 procamburus
34	3	5.2	7	2 O9YIO9	O9YIO9 human adeno
35	3	5.2	7	2 O9YIRO	O9YIRO human adeno
36	3	5.2	7	2 O9YVE3	O9YVE3 human adeno
37	3	5.2	8	1 AKHG_GRYBI	P67785 gryllus bim
38	3	5.2	8	1 ARH_ROWMI	P67786 romalea mic
39	3	5.2	8	1 CAD1_ENTPA	P13268 entrococcu
40	3	5.2	8	1 FARI_PANRE	P41872 panagrellus
41	3	5.2	8	1 FARI_PENNO	P83316 penaeus mon
42	3	5.2	8	1 PAR2_MACRS	P83375 macrobrachi
43	3	5.2	8	1 PAR3_HOMAM	P41486 homarus ame
44	3	5.2	8	1 FAR4_HOMAM	P41487 homarus ame
45	3	5.2	8	1 HTFI_PERAM	P04548 periplaneta

## ALIGNMENTS

RESULT 1	FLRN ANTEL	STANDARD	PRT	4 AA
AC	P58707			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Antio-RNamide.			
OS	Antiopeura elegantissima (Sea anemone).			
OC	Eukaryota; Metazoa; Chordata; Anthozoa; Zoantharia; Actinoptera;			
OC	Nymphaeae; Actinoptera; Anthoptera.			
OX	NCBI_TaxID=6110;			
RN	[1]			
RP	SEQUENCE, AND MASS SPECTROMETRY.			
RX	MEDLINE=90319122; PubMed=1973541;			
RA	Grimmelikhuizen C.J.P., Rinehart K.L. Jr., Jacob E., Graff D.,			
RA	Reinisch H.K., Nothacker H.-P., Staley A.L.,			
RT	"Isolation of L-3-phenylacetyl-Leu-Arg-Asn-NH2 (Antio-RNamide), a sea			
RT	anemone neuropeptide containing an unusual amino-terminal blocking			
RT	group."			
RL	Proc. Natl. Acad. Sci. U.S.A. 87:5410-5414(1990).			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- TISSUE SPECIFICITY: Neuron specific.			
CC	-1- MASS SPECTROMETRY: MW=549.3; METHOD=FPAB; RANGE=1-4; NOTE=Ref.1.			
CC	PIR: A35779; A35779.			
KW	Amidation; Direct protein sequencing; Neuropeptide.			
FT	MOD_RES 1 4 1 3-phenylacetic acid.			
FT	MOD_RES 4 4 4 Asparagine amide.			
SQ	SEQUENCE 4 AA; 549 MW; 64540729A0000000 CRC64;			
Query Match	6.9%; Score 4; DB 1; Length 4;			
Best Local Similarity	100.0%; Pred. No. 1.6e+06;			
Matches	4; Conservative			
	Mismatches 0; Indels 0; Gaps 0;			
QY	29 FLRN 32			
Db	1 FLRN 4			
RESULT 2				
ID	P92210	PRELIMINARY;	PRT;	7 AA.
AC	P92210;			
DT	01-MAY-1997 (TREMBLrel. 03, Created)			
DT	01-MAY-1997 (TREMBLrel. 03, Last sequence update)			
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DR	Ribosomal protein l1 (Fragment).			
GN	Name=rp81;			
OS	Agropyron cristatum (Crested wheatgrass).			
OC	Chloroplast.			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;			
OC	Triticaceae; Agropyron.			
OX	NCBI_TaxID=4593;			

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaves;
RX MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpcv.1996.0389;
RA Petersen G., Seberg O.;
RT "Phylogenetic analysis of the Trilicaceae (Poaceae) based on rpoA
sequence data.";
RL Mol. Phylogenet. Evol. 7:217-230(1997).
DR EMBL; 277771; CAB01391.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW Chloroplast; Ribosomal protein.
FT NON TER 1
SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match
Best Local Similarity 100.0%; Score 4; DB 2; Length 7;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKRR 46
DB 3 KKRR 6

RESULT 3
P92214 PRELIMINARY; PRT; 7 AA.
ID P92214;
AC P92214;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Ribosomal protein 11 (Fragment).
GN Name=rp511;
OS Amblyopyrum multicum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Trilicaceae; Amblyopyrum.
OX NCBI_TaxID=4595;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaves;
RX MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpcv.1996.0389;
RA Petersen G., Seberg O.;
RT "Phylogenetic analysis of the Trilicaceae (Poaceae) based on rpoA
sequence data.";
RL Mol. Phylogenet. Evol. 7:217-230(1997).
DR EMBL; 277756; CAB01346.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW Chloroplast; Ribosomal protein.
FT NON TER 1
SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match
Best Local Similarity 100.0%; Score 4; DB 2; Length 7;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKRR 46
DB 3 KKRR 6

RESULT 4
P92218 PRELIMINARY; PRT; 7 AA.
ID P92218;
AC P92218;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Ribosomal protein 11 (Fragment).
GN Name=rp511;
OS Austroalopyrum retrofractum.

```

```

OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Trilicaceae; Austroalopyrum.
OX NCBI_TaxID=4597;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaves;
RX MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpcv.1996.0389;
RA Petersen G., Seberg O.;
RT "Phylogenetic analysis of the Trilicaceae (Poaceae) based on rpoA
sequence data.";
RL Mol. Phylogenet. Evol. 7:217-230(1997).
DR EMBL; 277767; CAB01379.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW Chloroplast; Ribosomal protein.
FT NON TER 1
SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match
Best Local Similarity 100.0%; Score 4; DB 2; Length 7;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKRR 46
DB 3 KKRR 6

RESULT 5
P92221 PRELIMINARY; PRT; 7 AA.
ID P92221;
AC P92221;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Ribosomal protein 11 (Fragment).
GN Name=rp511;
OS Bromus inermis (Smooth brome grass).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Bromae; Bromus.
OX NCBI_TaxID=15371;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaves;
RX MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpcv.1996.0389;
RA Petersen G., Seberg O.;
RT "Phylogenetic analysis of the Trilicaceae (Poaceae) based on rpoA
sequence data.";
RL Mol. Phylogenet. Evol. 7:217-230(1997).
DR EMBL; 277759; CAB01355.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW Chloroplast; Ribosomal protein.
FT NON TER 1
SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match
Best Local Similarity 100.0%; Score 4; DB 2; Length 7;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKRR 46
DB 3 KKRR 6

RESULT 6
P92226 PRELIMINARY; PRT; 7 AA.
ID P92226;
AC P92226;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)

```

DT 01-MAY-1997 (TREMBlrel. 03, last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)  
DE Ribosomal protein 11 (Fragment).  
GN Name=rp51;  
OS Crithopsis delileana.  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Trilicaceae; Crithopsis.  
OX NCBI\_TaxId=37674;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=leaves;  
RX MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpcv.1996.0389;  
RA Petersen G., Seberg O.;  
RT "Phylogenetic analysis of the Trilicaceae (Poaceae) based on rpoA  
sequence data.";  
RL Mol. Phylogenet. Evol. 7:217-230(1997).  
DR EMBL; 277751; CAB01331.1; -;  
DR GO; GO:0009507; C:chloroplast; IEA.  
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
KW Chloroplast; Ribosomal protein.  
FT NON TER 1  
SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match 6.9%; Score 4; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 KKR 46  
Db 3 KKR 6

RESULT 7  
P92372 PRELIMINARY; PRT; 7 AA.  
ID P92372;  
AC P92372;  
DT 01-MAY-1997 (TREMBlrel. 03, Created)  
DT 01-MAY-1997 (TREMBlrel. 03, last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)  
DE Ribosomal protein 11 (Fragment).  
GN Name=rp51;  
OS Haynaldia villosa.  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Trilicaceae; Haynaldia.  
OX NCBI\_TaxId=40247;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=leaves;  
RX MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpcv.1996.0389;  
RA Petersen G., Seberg O.;  
RT "Phylogenetic analysis of the Trilicaceae (Poaceae) based on rpoA  
sequence data.";  
RL Mol. Phylogenet. Evol. 7:217-230(1997).  
DR EMBL; 277741; CAB01301.1; -;  
DR GO; GO:0009507; C:chloroplast; IEA.  
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
KW Chloroplast; Ribosomal protein.  
FT NON TER 1  
SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match 6.9%; Score 4; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 KKR 46  
Db 3 KKR 6

RESULT 8  
P92381 PRELIMINARY; PRT; 7 AA.  
ID P92381;  
AC P92381;  
DT 01-MAY-1997 (TREMBlrel. 03, Created)  
DT 01-MAY-1997 (TREMBlrel. 03, last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)  
DE Ribosomal protein 11 (Fragment).  
GN Name=rp51;  
OS Hordeum brachyantherum.  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Trilicaceae; Hordeum.  
OX NCBI\_TaxId=52712;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=leaves;  
RX MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpcv.1996.0389;  
RA Petersen G., Seberg O.;  
RT "Phylogenetic analysis of the Trilicaceae (Poaceae) based on rpoA  
sequence data.";  
RL Mol. Phylogenet. Evol. 7:217-230(1997).  
DR EMBL; 277761; CAB01361.1; -;  
DR GO; GO:0009507; C:chloroplast; IEA.  
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
KW Chloroplast; Ribosomal protein.  
FT NON TER 1  
SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match 6.9%; Score 4; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 KKR 46  
Db 3 KKR 6

RESULT 9  
P92385 PRELIMINARY; PRT; 7 AA.  
ID P92385;  
AC P92385;  
DT 01-MAY-1997 (TREMBlrel. 03, Created)  
DT 01-MAY-1997 (TREMBlrel. 03, last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)  
DE Ribosomal protein 11 (Fragment).  
GN Name=rp51;  
OS Hordeum marinum subsp. gussoneanum.  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Trilicaceae; Hordeum.  
OX NCBI\_TaxId=98114;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=leaves;  
RX MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpcv.1996.0389;  
RA Petersen G., Seberg O.;  
RT "Phylogenetic analysis of the Trilicaceae (Poaceae) based on rpoA  
sequence data.";  
RL Mol. Phylogenet. Evol. 7:217-230(1997).  
DR EMBL; 277763; CAB01367.1; -;  
DR GO; GO:0009507; C:chloroplast; IEA.  
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
KW Chloroplast; Ribosomal protein.  
FT NON TER 1  
SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match 6.9%; Score 4; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKR 46  
 Db 3 KKR 6

## RESULT 10

P23387 PRELIMINARY; PRT; 7 AA.  
 AC P23387; 01-MAY-1997 (TRENBLREL. 03, Created)  
 DT 01-MAY-1997 (TRENBLREL. 03, Last sequence update)  
 DE 01-JUN-2003 (TRENBLREL. 24, Last annotation update)  
 GN Ribosomal protein 11 (Fragment).  
 OS Name=rp11;  
 OS Heteranthellium piliiferum.  
 OC Bacteriophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticeae; Heteranthellium.  
 OK NCBI\_TaxID=37678;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=leaves;  
 RX MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpcv.1996.0389;  
 RA Petersen G., Seberg O.;  
 RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA  
 sequence data.";  
 RL Mol. Phylogenet. Evol. 7:217-230(1997).  
 DR EMBL; Z77748; CAB01322.1; -;  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 KW Chloroplast; Ribosomal protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match 6.9%; Score 4; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKR 46  
 Db 3 KKR 6

## RESULT 11

P23390 PRELIMINARY; PRT; 7 AA.  
 AC P23390; 01-MAY-1997 (TRENBLREL. 03, Created)  
 DT 01-MAY-1997 (TRENBLREL. 03, Last sequence update)  
 DE 01-JUN-2003 (TRENBLREL. 24, Last annotation update)  
 GN Ribosomal protein 11 (Fragment).  
 OS Name=rp11;  
 OS Heteranthellium piliiferum.  
 OC Chloroplast.  
 OC Bacteriophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticeae; Heteranthellium.  
 OK NCBI\_TaxID=37679;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=leaves;  
 RX MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpcv.1996.0389;  
 RA Petersen G., Seberg O.;  
 RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA  
 sequence data.";  
 RL Mol. Phylogenet. Evol. 7:217-230(1997).  
 DR EMBL; Z77750; CAB01328.1; -;  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 KW Chloroplast; Ribosomal protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match 6.9%; Score 4; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKR 46  
 Db 3 KKR 6

## RESULT 12

P23393 PRELIMINARY; PRT; 7 AA.  
 AC P23393; 01-MAY-1997 (TRENBLREL. 03, Created)  
 DT 01-MAY-1997 (TRENBLREL. 03, Last sequence update)  
 DE 01-JUN-2003 (TRENBLREL. 24, Last annotation update)  
 GN Ribosomal protein 11 (Fragment).  
 OS Name=rp11;  
 OS Hordeum vulgare (Barley).  
 OC Chloroplast.  
 OC Bacteriophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticeae; Hordeum.  
 OK NCBI\_TaxID=4513;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=leaves;  
 RX MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpcv.1996.0389;  
 RA Petersen G., Seberg O.;  
 RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA  
 sequence data.";  
 RL Mol. Phylogenet. Evol. 7:217-230(1997).  
 DR EMBL; Z77764; CAB01370.1; -;  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 KW Chloroplast; Ribosomal protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match 6.9%; Score 4; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKR 46  
 Db 3 KKR 6

## RESULT 13

P2403 PRELIMINARY; PRT; 7 AA.  
 AC P2403; 01-MAY-1997 (TRENBLREL. 03, Created)  
 DT 01-MAY-1997 (TRENBLREL. 03, Last sequence update)  
 DE 01-JUN-2003 (TRENBLREL. 24, Last annotation update)  
 GN Ribosomal protein 11 (Fragment).  
 OS Name=rp11;  
 OS Lophopyrum elongatum (Tall wheatgrass) (Argopyrum elongatum).  
 OC Chloroplast.  
 OC Bacteriophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticeae; Lophopyrum.  
 OK NCBI\_TaxID=4588;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=leaves;  
 RX MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpcv.1996.0389;  
 RA Petersen G., Seberg O.;  
 RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA  
 sequence data.";  
 RL Mol. Phylogenet. Evol. 7:217-230(1997).  
 DR EMBL; Z77743; CAB01307.1; -;



DR GO; GO:0009507; Chloroplast; IEA.  
 DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
 KW Chloroplast; Ribosomal protein.  
 FT NON TER 1  
 SQ SEQUENCE 7 AA; 894 MW; 6734140333277700 CRC64;

Query Match 6.9%; Score 4; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKR 46  
 Db 3 KKR 6

## RESULT 14

P92421 PRELIMINARY; PRT; 7 AA.  
 AC P92421; P92419;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Ribosomal protein 11 (Fragment).  
 GN Name=rp11;  
 OS Psathyrostachys fragilis.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticeae; Psathyrostachys.  
 OX NCBI\_TaxID=37729;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=leaves;  
 RX MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpcv.1996.0389;  
 RA Petersen G., Seberg O.;  
 RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA  
 RT sequence data."  
 RL Mol. Phylogenet. Evol. 7:217-230(1997).  
 DR EMBL; Z77753; CAB01337.1; -  
 DR GO; GO:0009507; Chloroplast; IEA.  
 DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
 KW Chloroplast; Ribosomal protein.  
 FT NON TER 1  
 SQ SEQUENCE 7 AA; 894 MW; 6734140333277700 CRC64;

Query Match 6.9%; Score 4; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKR 46  
 Db 3 KKR 6

## RESULT 15

P92425 PRELIMINARY; PRT; 7 AA.  
 AC P92425;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Ribosomal protein 11 (Fragment).  
 GN Name=rp11;  
 OS Pseudoroegneria spicata (Bluebunch wheatgrass) (Agropyron spicatum).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticeae; Pseudoroegneria.  
 OX NCBI\_TaxID=4604;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=leaves;  
 RX MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpcv.1996.0389;

RA Petersen G., Seberg O.;  
 RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA  
 RT sequence data."  
 RL Mol. Phylogenet. Evol. 7:217-230(1997).  
 DR EMBL; Z77753; CAB01337.1; -  
 DR GO; GO:0009507; Chloroplast; IEA.  
 DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
 KW Chloroplast; Ribosomal protein.  
 FT NON TER 1  
 SQ SEQUENCE 7 AA; 894 MW; 6734140333277700 CRC64;

Query Match 6.9%; Score 4; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKR 46  
 Db 3 KKR 6

Search completed: October 16, 2005, 05:59:29  
 Job time : 119 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 16, 2005, 05:51:50 ; Search time 51 Seconds

(without alignments)  
109.423 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 58

Sequence: 1 MQMFPSPLEFFFLQLKQSS.....RYGKKRRATRFMDPRRGTP 58

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1102

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	6.9	4	A35779	neuropeptide Antho
2	4	6.9	8	PC4131	hypothetical prote
3	4	6.9	10	A56633	neomysuppressin -
4	4	6.9	10	A32543	cardioexcitatory n
5	3	5.2	5	PT0525	T-cell receptor be
6	3	5.2	5	PT0577	T-cell receptor be
7	3	5.2	5	PT0700	T-cell receptor be
8	3	5.2	5	PT0565	T-cell receptor be
9	3	5.2	6	A11490	pyruvate kinase (B
10	3	5.2	6	PT0518	T-cell receptor be
11	3	5.2	7	S78024	ribosomal protein
12	3	5.2	7	I48086	DNA topoisomerase
13	3	5.2	7	A61081	tryptophyltin, bas
14	3	5.2	8	A44960	neuropeptide Ded-C
15	3	5.2	8	A61348	red pigment-concen
16	3	5.2	8	S08995	hypertrehalosemic
17	3	5.2	8	A49823	adipokinetic hormo
18	3	5.2	8	A28004	hypertrehalosemic
19	3	5.2	8	A43976	hypertrehalosemic
20	3	5.2	8	B43976	hypertrehalosemic
21	3	5.2	8	T10077	hypothetical prote
22	3	5.2	8	S19288	acylase - Kluverera
23	3	5.2	8	S21288	lectin - potato (f
24	3	5.2	8	A05169	neuropeptide M-I -
25	3	5.2	8	JS0316	leucokinin VI - Ma
26	3	5.2	8	S71919	alcohol dehydrogen
27	3	5.2	9	A61102	parathyroid hormon
28	3	5.2	9	PT0326	Ig heavy chain CRD
29	3	5.2	9	S19523	orf AB protein - S

30	3	5.2	9	A31576	xylose isomerase (
31	3	5.2	9	PC7076	spectrin alpha cha
32	3	5.2	10	A60624	angiotensin I - Ja
33	3	5.2	10	A60410	beta-neoendorphin
34	3	5.2	10	S08997	hypertrehalosemic
35	3	5.2	10	A60421	hypertrehalosemic
36	3	5.2	10	S08998	hypertrehalosemic
37	3	5.2	10	A26381	hypertrehalosemic
38	3	5.2	10	UN0440	peptide-N4-(N-acet
39	3	5.2	10	PO0753	beta-fructofuranos
40	3	5.2	10	A61622	vitellinogenin, 190k
41	3	5.2	10	C39111	Ig heavy chain C r
42	3	5.2	10	S10785	enamelin, 22k - bo
43	3	5.2	10	S70722	65.4k GTP-binding
44	3	5.2	10	PT0084	protein OA60021 -
45	3	5.2	10	B59272	peptide-N4-(N-acet

#### ALIGNMENTS

RESULT 1  
A35779  
neuropeptide Antho-RNamide - sea anemone (Anthopleura elegantissima)  
C/Species: Anthopleura elegantissima  
C/Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004  
C/Accession: A35779  
R/Girimeallikhuizen, C.U.P., Rinehart, K.L.; Jacob, E.; Graff, D.; Reinscheid, R.K.; Noth  
Proc. Natl. Acad. Sci. U.S.A. 87, 5410-5414, 1990  
A/Title: Isolation of L-3-phenylacetyl-L-tyr-Arg-Asn-NH2 (Antho-RNamide), a sea anemone ne  
A/Reference number: A35779; MUID:90319122; PMID:1973541  
A/Accession: A35779  
A/Molecule type: protein  
A/Residues: 1-4 <GRI>  
A/Cross-references: UNIPROT:P58707  
C/Comment: The L-3-phenylacetyl blocking group probably arises from an amino-terminal ph  
C/Keywords: amidated carboxyl end; neuropeptide; phenylacetylation  
P/1/Modified site: L-3-phenylacetic acid (Phe) #status experimental  
F/4/Modified site: amidated carboxyl end (Asn) #status experimental

Query Match 6.9%; Score 4; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 FLRN 32  
Db 1 FLRN 4

RESULT 2  
PC4131  
hypothetical protein 8 [imported] - Pseudomonas aeruginosa (fragment)  
C/Species: Pseudomonas aeruginosa  
C/Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C/Accession: PC4131

R/Kawasaki, S.; Arai, H.; Igatahshi, Y.; Kodama, T.  
Gene 167, 87-91, 1995  
A/Title: Sequencing and characterization of the downstream region of the genes encoding  
Y for biosynthesis of heme d1.

A/Reference number: JC4552; MUID:96144254; PMID:8566817  
A/Accession: PC4131  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-8 <KAW>

A/Cross-references: UNIPROT:P95412; DDBJ:D50473; NID:g1217594  
A/Note: this ORF is not annotated in GenBank entry PSENIRC, release 113.0

Query Match 6.9%; Score 4; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 SRR 33  
|||||

Db 5 SRRL 8

RESULT 3  
A:Accession: A56633  
neomysuppressin - flesh fly (Sarcophaga bullata)  
N:Alternate names: Neb-MS  
C:Species: Sarcophaga bullata  
C:Date: 21-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004  
C:Accession: A56633  
R:Feeney, A.; Schoofs, L.; Proost, P.; Van Damme, J.; Buedts, H.; De Loof, A.  
Comp. Biochem. Physiol. C 102, 239-245, 1992  
A:Title: Isolation, primary structure and synthesis of neomysuppressin, a myoinhibiting  
A:Reference number: A56633; MUID:93047886; PMID:1358537  
A:Accession: A56633  
A:Molecule type: protein  
A:Residues: 1-10 <RON>  
A:Cross-references: UNIPROT:P61850  
A:Experimental source: head  
A>Note: sequence extracted from NCBI backbone (NCBIP:119072)  
C:Keywords: amidated carboxyl end; neuropeptide  
F:10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match  
Best Local Similarity 100.0%; Score 4; DB 2; Length 10;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 VFLR 31  
Db 6 VFLR 9

RESULT 4  
A:Accession: A32543  
cardioexcitatory neuropeptide - desert locust  
C:Species: Schistocerca gregaria (desert locust)  
C:Date: 20-Dec-1989 #sequence\_revision 20-Dec-1989 #text\_change 09-Jul-2004  
C:Accession: A32543  
R:Robb, S.; Packman, L.C.; Evans, P.D.  
Biochem. Biophys. Res. Commun. 160, 850-856, 1989  
A:Title: Isolation, primary structure and bioactivity of Schistocerca gregaria, a FMRF-amide  
A:Reference number: A32543; MUID:89246543; PMID:2719702  
A:Accession: A32543  
A:Status: Preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <ROB>  
A:Cross-references: UNIPROT:P38553  
C:Superfamily: unassigned animal peptides  
C:Keywords: amidated carboxyl end; neuropeptide  
F:10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match  
Best Local Similarity 100.0%; Score 4; DB 2; Length 10;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 VFLR 31  
Db 6 VFLR 9

RESULT 5  
PT0525  
T-cell receptor beta chain V-D-J region (100-4J) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0525  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601; PMID:1711558  
A:Accession: PT0525  
A:Status: translation not shown  
A:Molecule type: mRNA

A:Residues: 1-5 <FEB>  
A:Experimental source: adult thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match  
Best Local Similarity 100.0%; Score 3; DB 2; Length 5;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 RGT 57  
Db 3 RGT 5

RESULT 6  
PT0577  
T-cell receptor beta chain V-D-J region (141-1BC) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0577; PT0574  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601; PMID:1711558  
A:Accession: PT0577  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-5 <FEB>  
A:Experimental source: day 19 fetal thymus, strain BALB/c, clone 141-1BC  
A:Accession: PT0574  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-5 <FEB>  
A:Experimental source: day 19 fetal thymus, strain BALB/c, clone 141-1Q  
C:Keywords: T-cell receptor

Query Match  
Best Local Similarity 100.0%; Score 3; DB 2; Length 5;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 SSR 21  
Db 2 SSR 4

RESULT 7  
PT0700  
T-cell receptor beta chain V-D-J region (161-2A) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0700  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601; PMID:1711558  
A:Accession: PT0700  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-5 <FEB>  
A:Experimental source: newborn thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match  
Best Local Similarity 100.0%; Score 3; DB 2; Length 5;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 SSR 21  
Db 2 SSR 4

RESULT 8  
PT0565  
T-cell receptor beta chain V-D-J region (141-1CF) - mouse (fragment)

C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 09-Jul-2004  
C:Accession: PT0565  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601; PMID:1711558  
C:Accession: PT0565  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-5 <FE8>  
A:Cross-references: UNIPROT:Q8B2Q7  
A:Experimental source: day 19 fetal thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 5.2%; Score 3; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 SSR 21  
DB 2 SSR 4

RESULT 9  
A11490  
pyruvate kinase (EC 2.7.1.40) - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 03-Mar-1995  
C:Accession: A11490  
R:Hjeltnes, G.; Andersson, J.; Edlund, B.; Engstrom, L.  
Biochem. Biophys. Res. Commun. 61, 559-563, 1974  
A:Title: Amino acid sequence of a (32-P)phosphopeptide from pig liver pyruvate kinase ph  
A:Reference number: A11490; MUID:75127438; PMID:4375589  
A:Accession: A11490  
A:Molecule type: protein  
A:Residues: 1-6 <HUB>  
A:Experimental source: liver  
C:Keywords: glycolysis; phosphotransferase

Query Match 5.2%; Score 3; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 RRA 47  
DB 2 RRA 4

RESULT 10  
PT0518  
T-cell receptor beta chain V-D-J region (100-4B) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0518  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601; PMID:1711558  
A:Accession: PT0518  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-6 <FE8>  
A:Experimental source: adult thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 5.2%; Score 3; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 RAT 48  
DB 3 RAT 5

## RESULT 11

578024  
ribosomal protein Yml21, mitochondrial - yeast (Saccharomyces cerevisiae) (fragment)  
C:Species: Saccharomyces cerevisiae  
C:Date: 09-Oct-1997 #sequence\_revision 24-Oct-1997 #text\_change 14-Nov-1997  
C:Accession: 578024  
R:Kitakawa, M.; Grack, H.R.; Grohmann, L.; Goldschmidt-Reisin, S.; Herfurth, E.; Wiltma  
Bur. J. Biochem. 245, 449-456, 1997  
A:Title: Identification and characterization of the genes for mitochondrial ribosomal pr  
A:Reference number: 578018; MUID:97296414; PMID:9151978  
A:Accession: 578024  
A:Molecule type: protein  
A:Residues: 1-7 <KIT>  
C:Genetics:  
A:Genome: nuclear  
C:Keywords: mitochondrion; protein biosynthesis; ribosome

Query Match 5.2%; Score 3; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LUK 17  
DB 5 LUK 7

RESULT 12  
148086  
DNA topoisomerase II alpha - Chinese hamster (fragment)  
C:Species: Crictulus griseus (Chinese hamster)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
C:Accession: 148086  
R:Ng, S.W.; Eder, J.P.; Schnipper, L.B.; Chan, V.T.W.  
J. Biol. Chem. 270, 25850-25858, 1995  
A:Title: Molecular cloning and characterization of the promoter for the Chinese hamster  
A:Reference number: 148086; MUID:96029684; PMID:7592770  
A:Accession: 148086  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-7 <RES>  
A:Cross-references: EMBL:U34196; NID:G1041231; PIDN:AC52315.1; PID:G1041232

Query Match 5.2%; Score 3; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SPL 9  
DB 4 SPL 6

RESULT 13  
A61081  
tryptophyltin, basic - Rohde's leaf frog  
C:Species: Phyllomedusa rohdei (Rohde's leaf frog)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 18-Aug-2000  
C:Accession: A61081  
R:Montecucchi, P.C.; Vincenti, M.; Lazzarini, A.M.; Ruesconi, L.; Erspamer, V.  
Int. J. Pept. Protein Res. 33, 391-395, 1989  
A:Title: Isolation, structure determination and synthesis of a novel tryptophan-containi  
A:Reference number: A61081  
A:Accession: A61081  
A:Molecule type: protein  
A:Residues: 1-7 <MON>  
C:Comment: The biological activity of this peptide was not determined.  
C:Superfamily: unassigned animal peptides  
C:Keywords: amidated carboxyl end; hydroxyproline; skin  
F:3/Modified site: 4-hydroxyproline (Pro) #status experimental  
F:7/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 7;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 PPS 7  
|||  
Db 2 PPS 4

## RESULT 14

A4960

neuropeptide led-CC-I - Colorado potato beetle

C:Species: Lepidocnara decemlineata (Colorado potato beetle)

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004

C:Accession: A4960

R:Gaede, G.; Kellner, R.

Peptides 10, 1287-1289, 1989

A:Title: The metabolic neuropeptides of the corpus cardiacum from the potato beetle and

A:Reference number: A4960; PMID:90160053; PMID:2576128

A:Accession: A4960

A:Molecule type: protein

A:Residues: 1-8 &lt;GAB&gt;

A:Cross-references: UNIPROT:P04548

C:Superfamily: adipokinetic hormone

C:Keywords: blocked carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:8/Modified site: blocked carboxyl end (Trp) (probably amidated) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 NFS 34  
|||  
Db 3 NFS 5

## RESULT 15

A61348

red pigment-concentrating hormone - northern shrimp

N:Alternate names: blanching hormone

C:Species: Pandalus borealis (northern shrimp)

C:Date: 02-Aug-1994 #sequence\_revision 05-Aug-1994 #text\_change 09-Jul-2004

C:Accession: A61348; S07139

R:Fernlund, P.; Josefsson, L.

Science 177, 173-175, 1972

A:Title: Crustacean color-change hormone: amino acid sequence and chemical synthesis.

A:Reference number: A61348; PMID:72228738; PMID:5041363

A:Accession: A61348

A:Molecule type: protein

A:Residues: 1-8 &lt;FERL&gt;

A:Cross-references: UNIPROT:P08939

R:Fernlund, P.

Biochim. Biophys. Acta 371, 304-311, 1974

A:Title: Structure of the red-pigment-concentrating hormone of the shrimp, Pandalus bore

A:Reference number: S07139; PMID:75054965; PMID:4433569

A:Accession: S07139

A:Molecule type: protein

A:Residues: 'E', 2-8 &lt;FER2&gt;

A:Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have

C:Comment: This neuropeptide, isolated from the eyestalks of the shrimp and active in pi

zed pigment-containing cells.

C:Superfamily: adipokinetic hormone

C:Keywords: amidated carboxyl end; blocked amino end; hormone; neuropeptide; pyroglutamic

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 NFS 34  
|||  
Db 3 NFS 5

Search completed: October 16, 2005, 06:00:31  
Job time : 52 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 16, 2005, 05:44:34 ; Search time 117 Seconds  
(without alignments)  
191.727 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 58  
Sequence: 1 MQMFPSPPLFFFLQLKQSS.....RYIGKKRATRFWDPRGTP 58

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 465227

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_1Dec04:\*

1: geneeqp1980s:\*\n2: geneeqp1990s:\*\n3: geneeqp2000s:\*\n4: geneeqp2001s:\*\n5: geneeqp2002s:\*\n6: geneeqp2003as:\*\n7: geneeqp2003bs:\*\n8: geneeqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	10.3	9	8	ADH89809 Cell pene
2	5	8.6	5	2	AAW51344 Peptide #
3	5	8.6	5	2	AAW51338 Peptide #
4	5	8.6	5	3	AAy49996 Natural b
5	5	8.6	5	4	AAE12536 AP70 abe
6	5	8.6	5	4	AAH80747 hK2 cleav
7	5	8.6	5	6	ABG73485 Natural b
8	5	8.6	7	4	AAH80706 Human gla
9	5	8.6	7	4	AAH80669 Human gla
10	5	8.6	7	4	AAH85856 Prostata-
11	5	8.6	7	7	ADB79449 Parapoxvi
12	5	8.6	7	8	ADN08587 Ceramic t
13	5	8.6	7	8	ADP74892 Parapoxvi
14	5	8.6	8	3	AAH06282 prosKI-1
15	5	8.6	9	2	AAH59128 Peptide f
16	5	8.6	9	2	AAy38039 Hepatitis
17	5	8.6	9	2	AAy38106 Hepatitis
18	5	8.6	9	2	AAW44612 Anti-fung
19	5	8.6	9	2	AAW43777 Bacterici
20	5	8.6	9	2	AAy45608 Immunogen
21	5	8.6	9	2	AAy45674 Immunogen
22	5	8.6	9	2	AAy46768 Immunogen
23	5	8.6	9	2	AAy00589 Anti-funga
24	5	8.6	9	4	AAH65513 Anti-fung
25	5	8.6	9	4	AAU00796 HIV-1 Tat

26	5	8.6	9	5	ABH08361 Synthetic
27	5	8.6	9	5	AAO15791 Human imm
28	5	8.6	9	5	ABP54737 HIV-1 tat
29	5	8.6	9	6	AAE35004 Mouse imm
30	5	8.6	9	7	ABH82213 Human ant
31	5	8.6	9	7	ADK65211 Human PIG
32	5	8.6	9	8	ADH97590 Immunogen
33	5	8.6	9	8	ADH98910 Cell pene
34	5	8.6	9	8	ADH44385 INK4a imm
35	5	8.6	9	8	ADH44376 INK4a imm
36	5	8.6	9	8	ADK84035 Human 191
37	5	8.6	9	8	ADK84519 Human 191
38	5	8.6	9	8	ADK85026 Human 191
39	5	8.6	9	8	ADK85074 Human 191
40	5	8.6	9	8	ADK87550 Human 191
41	5	8.6	9	8	ADK88447 Human 191
42	5	8.6	9	8	ADK83950 Human 191
43	5	8.6	9	8	ADK88445 Human 191
44	5	8.6	9	8	ADK84560 Human 191
45	5	8.6	9	8	ADK85620 Human 191

#### ALIGNMENTS

RESULT 1	ADH89809	standard; peptide; 9 AA.
ID	ADH89809	standard; peptide; 9 AA.
AC	ADH89809;	
XX		
DT	22-APR-2004	(first entry)
XX		
DE	Cell penetrating peptide (CPP) identification method-related peptide 111.	
XX		
KW	cell-penetrating peptide; CPP; bulk property value Z-B; Z-E1; Z-E2; Z-E3; Z-E4; Z-E5; antidiabetic; neuroprotective; nootropic; antiparkinsonian; cardiac; cytosolic; tranquiliser; immunosuppressive; antidepressant; anticonvulsant; antiinflammatory; analgesic; neuroleptic; antihistaminological; antitumor; cell-penetrating; infectious disease; diabetes type I; diabetes type II; Alzheimer's disease; Parkinson's disease; cancer; prion disease; cardiovascular disease; signal transduction.	
KW		
OS	Unidentified.	
XX		
PN	WO2003106491-A2.	
XX		
PD	24-DEC-2003.	
XX		
PR	18-JUN-2003; 2003WO-IB003163.	
XX		
PR	18-JUN-2002; 2002SE-00001863.	
XX		
PR	25-JUN-2002; 2002US-0391788P.	
XX		
PA	(CEPE-) CEPEP AB.	
XX		
PI	Haellbink M, Pooga M, Metsis M, Kogerman P, Valkna A, Melkas A; Lindgren M, Graesslund A, Eriksson G, Oestensson CG, Buddina M; Zorzo M, Elmquist A, Soomets U, Lundberg P, Järver P, Saar K; El-Andalousi S, Kilik K, Langel U;	
PI		
DR	WPI; 2004-090832/09.	
XX		
PT	Predicting, designing, detecting, and/or verifying novel cell-penetrating peptide based on assessment of bulk property value of sequences of cell-penetrating peptide.	
XX		
PS	Disclosure; Page 31; 148pp; English.	
CC	This invention relates to a novel method of identifying, designing, detecting, and/or verifying novel cell-penetrating peptide (CPP) based on assessment of bulk property value Z-E of sequences of CPP comprising 5 or	

CC more individual average interval values Z-E1, Z-E2, Z-E3, Z-E4 and Z-E5,  
 CC where Z-E1, Z-E2, Z-E3, Z-E4 and Z-E5 are average values of the  
 CC respective descriptor values for the residues in the amino acid sequence.  
 CC The invention may be useful for the development of compounds with an  
 CC antidiabetic, neuroprotective, nootropic, antiparkinsonian, cardiant,  
 CC cyrostatic, tranquiliser, immunosuppressive, antidepressant,  
 CC anticonvulsant, antiinflammatory, analgesic, neuroleptic,  
 CC ophthalmological or anticancer activity as a stimulator of cell-  
 CC penetration. The method of the invention is useful for identifying a cell  
 CC -penetrating peptide or protein and/or a cell-penetrating fragment of a  
 CC peptide or protein. In addition, the invention may be useful for checking  
 CC cellular penetration properties of a peptide, for producing a cell-  
 CC penetrating and functional protein-mimicking peptide and for de novo  
 CC design and production of an artificial cell-penetrating and/or and  
 CC artificial cell-penetrating and functional protein-mimicking peptide.  
 CC Compositions developed within the scope of the present invention may be  
 CC useful for treating infectious diseases, diabetes type I, diabetes type  
 CC II, Alzheimer's disease, Parkinson's disease, cancer, prion disease,  
 CC cardiovascular disease or disorders resulting from perturbed signal  
 CC transduction. The method of the invention is fast, efficient and reliable  
 CC for identifying, detecting, designing CPPs and for screening cellular  
 CC uptake of a broad variety of CPPs in vitro and in vivo. The present  
 CC sequence is that of a peptide which is related to the invention.

CC Sequence 9 AA;

Query Match 10.3%; Score 6; DB 8; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FLOLK 17  
 Db 4 FLOLK 9

RESULT 2

AAW51344  
 ID AAW51344 standard; peptide; 5 AA.

AAW51344;

14-AUG-1998 (first entry)

Peptide #27 useful as modulator of beta-amyloid peptide aggregation.

Natural beta-amyloid peptide; aggregation; D-amino acid;  
 Alzheimer's disease; beta-amyloidosis.

Synthetic.

Homo sapiens.

Key Location/Qualifiers

Modified-site 1 /note= "D-form residue, N-terminal acetyl"

Misc-difference 2 /note= "D-form residue"

Misc-difference 3 /note= "D-form residue"

Misc-difference 4 /note= "D-form residue"

Modified-site 5 /note= "D-form residue, C-terminal amide"

WO9808868-A1.

05-MAR-1998.

27-AUG-1997; 97WO-US015166.

27-AUG-1996; 96US-00703675.

21-JUL-1997; 97US-00897342.

(PRAE-) PRAECTS PHARM INC.

XX Pindels MA, Gelfer MT, Musso G, Signer ER, Wakefield J;  
 PI Molinoux S, Chin J, Lee J, Kelley M, Komar-Panicucci S;  
 PI Arico-Muendel CC, Phillips K, Hayward NJ;  
 DR WPI, 1998-216936/19.

XX Peptide compounds which are preferably based on beta-amyloid peptide(s) -  
 PT are useful in treatment of disorders related to beta-amyloidosis,  
 PT especially Alzheimer's disease.

PS Claim 8; Page 80; 92pp; English.

XX The invention relates to peptides that modulate natural beta-amyloid  
 CC peptide aggregation. The modulators of the invention comprise a peptide  
 CC preferably based on a beta-amyloid peptide, that is comprised entirely of  
 CC D-amino acids. Preferably the peptide comprises 3-5 D-amino acid residues  
 CC and includes at least two D-amino acid residues independently selected  
 CC from the group consisting of D-leucine, D-phenylalanine and D-valine.  
 CC Preferred amino-terminal modifying groups include cyclic, heterocyclic,  
 CC polycyclic and branched alkyl groups. Preferred carboxy-terminal  
 CC modifying groups include an amide group, an alkyl amide group, an aryl  
 CC amide group or a hydroxy group. The peptides may be used to treat  
 CC disorders associated with beta-amyloidosis especially Alzheimer's  
 CC disease. They may also be used in methods for detecting the presence of  
 CC beta-amyloid peptides in biological samples. The present sequence  
 CC represents a specifically claimed peptide

Sequence 5 AA;

Query Match 8.6%; Score 5; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LPFFL 13  
 Db 1 LPFFL 5

RESULT 3

AAW51338  
 ID AAW51338 standard; peptide; 5 AA.

AAW51338;

14-AUG-1998 (first entry)

Peptide #21 useful as modulator of beta-amyloid peptide aggregation.

Natural beta-amyloid peptide; aggregation; D-amino acid;  
 Alzheimer's disease; beta-amyloidosis.

Synthetic.

Homo sapiens.

Key Location/Qualifiers

Modified-site 1 /note= "D-form residue"

Misc-difference 2 /note= "D-form residue"

Misc-difference 3 /note= "D-form residue"

Misc-difference 4 /note= "D-form residue"

Modified-site 5 /note= "D-form residue, C-terminal amide"

WO9808868-A1.

05-MAR-1998.

27-AUG-1997; 97WO-US015166.



PR 27-AUG-1996; 96US-00703675.  
 PR 21-JUL-1997; 97US-00897342.  
 XX  
 PA (PRAE-) PRAECIS PHARM INC.  
 PI FIndels MA, Gefter ML, Musso G, Signer ER, Wakefield J;  
 PI Mollineaux S, Chin J, Kelley M, Komar-Panicucci S;  
 PI Arico-Mwendel CC, Phillips K, Hayward NJ;  
 DR WPI; 1998-216936/19.  
 XX  
 PT Peptide compounds which are preferably based on beta-amyloid peptide (s) -  
 PT are useful in treatment of disorders related to beta-amyloidosis,  
 PT especially Alzheimer's disease.  
 PS Claim 8; Page 78; 92pp; English.  
 XX  
 PS The invention relates to peptides that modulate natural beta-amyloid  
 CC peptide aggregation. The modulators of the invention comprise a peptide  
 CC preferably based on a beta-amyloid peptide, that is comprised entirely of  
 CC D-amino acids. Preferably the peptide comprises 3-5 D-amino acid residues  
 CC and includes at least two D-amino acid residues independently selected  
 CC from the group consisting of D-leucine, D-phenylalanine and D-valine.  
 CC Preferred amino-terminal modifying groups include cyclic, heterocyclic,  
 CC polycyclic and branched alkyl groups. Preferred carboxy-terminal  
 CC modifying groups include an amide group, an alkyl amide group, an aryl  
 CC amide group or a hydroxy group. The peptides may be used to treat  
 CC disorders associated with beta-amyloidosis, especially Alzheimer's  
 CC disease. They may also be used in methods for detecting the presence of  
 CC beta-amyloid peptides in biological samples. The present sequence  
 CC represents a specifically claimed peptide  
 SQ Sequence 5 AA;  
 QY Query Match 8.6%; Score 5; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 9 LFFFL 13  
 1 LFFFL 5  
 RESULT 4  
 AAY49996  
 ID AAY49996 standard; peptide; 5 AA.  
 XX  
 AC AAY49996;  
 XX  
 DT 04-FEB-2000 (first entry)  
 XX  
 DE Natural beta amyloid peptide aggregation modulating compound #3.  
 XX  
 KM Beta amyloid peptide aggregation; modulation; D-amino acid; diagnosis;  
 KM retro-inverso isomer; amyloidogenic disease; Alzheimer's disease;  
 KM Down's syndrome; hereditary cerebral haemorrhage; HCHWA-D;  
 KM amyloidosis-Dutch-type.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 OS  
 FT Key Location/Qualifiers  
 FT Misc-difference 1.5 "D-form residues"  
 FT Modified-site 5 /note="amided"  
 FT  
 PN US5985242-A.  
 XX  
 PD 16-NOV-1999.  
 XX  
 PF 27-AUG-1997; 97US-00920162.  
 XX

PR 27-OCT-1995; 95US-00548998.  
 PR 14-MAR-1996; 96US-00616081.  
 PR 27-AUG-1996; 96US-00703675.  
 PR 21-JUL-1997; 97US-00897342.  
 XX  
 PA (PRAE-) PRAECIS PHARM INC.  
 PI Wakefield J, Mollineaux S, Signer ER, Kelley M, Komar-Panicucci S;  
 PI Musso G, Phillips K, Hayward NJ, Gefter ML, FIndels MA, Lee J;  
 PI Arico-Mwendel CC, Chin J;  
 DR WPI; 2000-022266/02.  
 XX  
 PT Compound comprising a peptidic structure, an amino-terminal modifying  
 PT group and a carboxy-terminal modifying group, useful for treating  
 PT Alzheimer's disease.  
 PS Claim 8; Col 63; 40pp; English.  
 XX  
 PS The present invention describes a compound of formula A-(Xaa)-B,  
 CC comprising a peptidic structure (Xaa), an amino-terminal modifying group  
 CC (A), and a carboxy-terminal modifying group (B). AAY49973 to AAY49994  
 CC represent specifically claimed examples of (Xaa). Also described is a  
 CC method for inhibiting aggregation of natural beta-amyloid peptides and  
 CC treating Alzheimer's disease, comprising contacting the amyloid peptides  
 CC with A-(Xaa)-B; and a method for detecting the presence of natural beta-  
 CC amyloid peptides in a biological sample by contacting with A-(Xaa)-B. The  
 CC compound is useful for treating Alzheimer's disease and for detecting the  
 CC presence of natural beta-amyloid peptides in a biological sample. The  
 CC compound can also be used prophylactically or therapeutically to treat  
 CC other clinical occurrences of beta-amyloid deposition, such as in Down's  
 CC syndrome individuals and in patients with hereditary cerebral haemorrhage  
 CC with amyloidosis-Dutch-type (HCHWA-D). The present sequence represents an  
 CC example of the compound from the present invention  
 SQ Sequence 5 AA;  
 QY Query Match 8.6%; Score 5; DB 3; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 9 LFFFL 13  
 1 LFFFL 5  
 RESULT 5  
 AAB12536  
 ID AAB12536 standard; peptide; 5 AA.  
 XX  
 AC AAB12536;  
 XX  
 DT 03-JAN-2002 (first entry)  
 XX  
 DE APP70 abeta peptide (residues 17-21) Inverso isomer mutant (V18F; A21L).  
 XX  
 KM Beta-amyloid peptide; beta-AP; Abeta peptide; therapy; beta-amyloidosis;  
 KM Alzheimer's disease; neuroprotective; beta-amyloid precursor protein;  
 KM APP-770; nootropic; mutant; mutlein.  
 XX  
 OS Unidentified.  
 OS  
 FT Key Location/Qualifiers  
 FT Misc-difference 1.5 "D-form residues"  
 FT Modified-site 1 /note="D-form residues"  
 FT Modified-site 1 /note="Optionally modified with 4-hydroxy benzoyl or  
 FT acetyl"  
 FT Misc-difference 2 /note="Wild-type Val substituted with Phe; corresponds  
 FT to position 18 of natural Abeta peptide"  
 FT Misc-difference 5 /note="Wild-type Ala substituted with Leu; corresponds



CC natural beta-amyloid peptides and for treating a disorder associated with  
 CC beta-amyloidosis, e.g. Alzheimer's disease. The compound is also useful  
 CC for detecting the presence or absence of natural beta-amyloid peptides in  
 CC a biological sample, by contacting the biological sample with the  
 CC compound in vitro, where the compound is labelled with a detectable  
 CC substance and detecting the compound bound to natural beta-amyloid  
 CC peptides to thus detect the presence or absence of natural beta-amyloid  
 CC peptides. The compound is preferably labelled with radioactive technetium  
 CC or radioactive iodine and is administered to the subject. The compound is  
 CC also useful for prophylactically or therapeutically treating other  
 CC clinical occurrences of beta-amyloid deposition such as in Down's  
 CC syndrome, individuals and in patients with hereditary cerebral haemorrhage  
 CC with amyloidosis-Dutch-type (HCHWA-D), and for therapeutically allowing  
 CC for the sequestration of beta-amyloid proteins at sites that do not lead  
 CC to neurological impairment. This sequence represents a natural beta-  
 CC amyloid peptide modulator compound of the invention

XX  
 SQ Sequence 5 AA;

Query Match 8.6%; Score 5; DB 6; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 LEFFL 13  
 DB 1 LEFFL 5

RESULT 8  
 AAB80706 AAB80706 standard; peptide; 7 AA.

XX AAB80706;

DT 02-MAY-2001 (first entry)

DE Human glandular kallikrein 2 substrate peptide #36.

XX Cleavage; kallikrein 2; HK2; prodrug.

XX Synthetic.

XX WO200109165-A2.

XX 08-FEB-2001.

PF 28-JUL-2000; 2000WO-US040496.

PR 29-JUL-1999; 99US-0146316P.

PA (UYGO ) UNIV JOHNS HOPKINS.

PI Denneade SR, Isaacs JT, Lilja H, Christensen SB;

DR WPI; 2001-191450/19.

PT New peptides containing cleavage sites specifically cleaved by human  
 PT kallikrein 2, useful for producing prodrugs which treat HK2-producing  
 PT cell proliferative disorders without exhibiting non-specific toxicity.

XX Example 8; Page 29; 38pp; English.

CC The present invention relates to a peptide comprising an amino acid  
 CC sequence having a cleavage site specific for an enzyme having a  
 CC proteolytic activity of human kallikrein 2 (HK2), and which is up to 20  
 CC amino acids in length. The invention is useful for producing a prodrug  
 CC which involves linking a drug which contains a primary amine to the  
 CC peptide, in which the linking of the peptide to the drug inhibits the  
 CC therapeutic activity of the drug

XX Sequence 7 AA;

Query Match 8.6%; Score 5; DB 4; Length 7;

Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 GKRR 46  
 DB 2 GKRR 6

RESULT 9

AAB80669 AAB80669 standard; peptide; 7 AA.

XX AAB80669;

DT 02-MAY-2001 (first entry)

DE Human glandular kallikrein 2 cleavage site peptide #30.

XX Cleavage; kallikrein 2; HK2; prodrug.

XX Synthetic.

XX WO200109165-A2.

XX 08-FEB-2001.

PF 28-JUL-2000; 2000WO-US040496.

PR 29-JUL-1999; 99US-0146316P.

PA (UYGO ) UNIV JOHNS HOPKINS.

PI Denneade SR, Isaacs JT, Lilja H, Christensen SB;

DR WPI; 2001-191450/19.

PT New peptides containing cleavage sites specifically cleaved by human  
 PT kallikrein 2, useful for producing prodrugs which treat HK2-producing  
 PT cell proliferative disorders without exhibiting non-specific toxicity.

XX Disclosure; Page 8; 38pp; English.

CC The present invention relates to a peptide comprising an amino acid  
 CC sequence having a cleavage site specific for an enzyme having a  
 CC proteolytic activity of human kallikrein 2 (HK2), and which is up to 20  
 CC amino acids in length. The invention is useful for producing a prodrug  
 CC which involves linking a drug which contains a primary amine to the  
 CC peptide, in which the linking of the peptide to the drug inhibits the  
 CC therapeutic activity of the drug

XX Sequence 7 AA;

Query Match 8.6%; Score 5; DB 4; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 GKRR 46  
 DB 2 GKRR 6

RESULT 10

AAB85856 AAB85856 standard; peptide; 7 AA.

XX AAB85856;

DT 08-MAY-2002 (first entry)

DE Prostate-specific antigen (PSA)-specific peptide #42.

XX Prostate-specific antigen; PSA; cytostatic; sesquiterpene-gamma-lactone;  
 KW cell proliferative disorder; malignant cell; non-malignant cell;

KM prostate cancer; thapsigargin.  
 XX Synthetic.  
 OS  
 XX WO200193861-A1.  
 PN  
 XX 13-DEC-2001.  
 PD  
 XX 28-JUL-2000; 2000WO-US020542.  
 PF  
 XX 07-JUN-2000; 2000US-00588822.  
 PR  
 XX (UYJO ) UNITV JOHNS HOPKINS.  
 PA  
 XX Isaacs JT, Lilja H, Denmeade SR, Christensen SB;  
 PI  
 XX WPI; 2002-216722/27.  
 DR  
 XX  
 PT New sesquiterpene-gamma-lactone comprising thapsigargin, is useful for  
 PS treating cell proliferative disorders.  
 XX  
 PS Disclosure; Page 18; 72pp; English.  
 CC The invention relates to a novel therapeutic sesquiterpene-gamma-lactone  
 CC comprising thapsigargin, and prostate antigen-specific (PSA) peptides.  
 CC These are used for treating cell proliferative disorders including  
 CC malignant as well as non-malignant cell populations, such as in prostate  
 CC cancer. The derivative can be linked to a variety of carrier moieties,  
 CC such as the peptides of the invention. The linkage converts the PSA-  
 CC derivative into an inactive prodrug. AAUS818-AAUS861 represent the PSA-  
 CC specific peptides of the invention  
 XX  
 SQ Sequence 7 AA;

Query Match 8.6%; Score 5; DB 5; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 GKKRR 46  
 |||||  
 Db 2 GKKRR 6

RESULT 11  
 ADB79449  
 ID ADB79449 strand; peptide; 7 AA.  
 XX  
 AC ADB79449;  
 XX  
 DT 04-DEC-2003 (first entry)  
 XX  
 DE Parapoxvirus ORF 12r N-terminal peptide.  
 XX  
 KM virulence; anti-HIV; hepatotropic; antiinflammatory; cytostatic;  
 KM vulnerability; antitubercular; antiallergic; dermatological; antidiabetic;  
 KM immunosuppressive; antineoplastic; antitubercular; thyromimetic;  
 KM procoagulant; amebicidal; antibacterial; gene therapy; virus;  
 KM viral infections; non-viral infections; proliferative disease;  
 KM inflammatory disease; allergic disease; autoimmune disease.  
 XX  
 OS Parapoxvirus.  
 XX  
 PN WO2003006654-A2.  
 PD  
 XX 23-JAN-2003.  
 PF  
 XX 12-JUN-2002; 2002WO-EP006440.  
 PR  
 XX 13-JUN-2001; 2001NZ-00512341.  
 PA  
 XX (FARB ) BAYER AG.  
 PI Weber O, Friederichs SM, Siegling A, Schlapp T, Mercer AA;

PI Fleming SB;  
 XX  
 DR WPI; 2003-221750/21.  
 XX  
 PT New polynucleotide and recombinant proteins of Parapoxvirus ovis, useful  
 PT for manufacturing a medicament for treating virus related disease, viral  
 PT infections, non-viral infections, proliferative disease or inflammatory  
 PT disease.  
 XX  
 PS Example 4; Page 33; 51pp; English.

CC The invention relates to a novel purified and isolated polynucleotide  
 CC (N1) of Parapoxvirus ovis (PPVO) comprising a nucleotide sequence (S1),  
 CC not defined in the specification, or its complementary sequence,  
 CC fragment or functional variant. A polynucleotide of the invention has  
 CC virulence, anti-HIV, hepatotropic, antiinflammatory, cytostatic,  
 CC vulnerability, antitubercular, antiallergic, dermatological, antidiabetic,  
 CC immunosuppressive, antineoplastic, antitubercular, thyromimetic,  
 CC procoagulant, amebicidal, and antibacterial activity. The polynucleotides  
 CC may have a use in gene therapy. The recombinant proteins encoded by the  
 CC polynucleotides, or recombinant viruses comprising a Vaccinia virus  
 CC genome and fragments of a PPVO genome are useful for manufacturing  
 CC pharmaceutical compositions for treating virus related disease (e.g.  
 CC hepatitis, papillomatosis, herpes virus infections, liver fibrosis, HIV  
 CC infections or influenza), viral infections, non-viral infections (e.g.  
 CC infections with mycobacteria, mycoplasma, amoeba or plasmodia),  
 CC proliferative disease (e.g. cancer, leukaemia, warts or other skin  
 CC neoplasms), inflammatory disease (e.g. Crohn's disease, COPD, asthma or  
 CC conditions related to healing of wounds), allergic disease, and/or  
 CC autoimmune diseases (systemic lupus erythematosus, Sjogren's disease,  
 CC Hashimoto's thyroiditis, rheumatoid arthritis or diabetes mellitus). The  
 CC present sequence is used in the exemplification of the invention.  
 XX  
 SQ Sequence 7 AA;

Query Match 8.6%; Score 5; DB 7; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 LQLLK 17  
 |||||  
 Db 2 LQLLK 6

RESULT 12  
 ADN08587  
 ID ADN08587 strand; peptide; 7 AA.  
 XX  
 AC ADN08587;  
 XX  
 DT 15-JUL-2004 (first entry)  
 XX  
 DE Ceramic tea strain binding peptide, SEQ ID No 152.  
 XX  
 KM tannin; polyphenolic; anthocyanin; tea; wine; stain; fabric; surface;  
 KM enzymatic; ceramic; hair; skin; detergent; cleaning; wine ageing;  
 KM fermentation; reducing; eliminating; astringency; conjugate; bleaching;  
 KM personal care product; cream; lotion; ointment.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200403482-A2.  
 PD  
 XX 22-APR-2004.  
 PF  
 XX 06-OCT-2003; 2003WO-US031776.  
 PR  
 XX 08-OCT-2002; 2002US-0417210P.  
 PA  
 XX (GENEV ) GENENCOR INT INC.  
 PI Murray CJ, Tijerina P, Van Gastel FUG;

DR WPI; 2004-340885/31.  
XX Novel binding peptide binding to tannin, anthocyanin and phenolic  
PT compounds, useful for delivering agent to target, and for modifying tea  
PT or wine stain on fabric or surface chosen from ceramic, glasses, wood,  
PT paper, skin, hair and plastic.  
XX Claim 1, SEQ ID NO 152; 38pp; English.  
XX  
CC The invention relates to a novel peptide which binds to tannin,  
CC polyphenolic or anthocyanin compounds. The binding peptide consists  
CC essentially of any one of 316 amino acid sequences of 5-13 amino acids in  
CC length or has at least 70% identity to one of the 316 amino acid  
CC sequences. The binding peptide is useful for delivering an agent to a  
CC target, which involves conjugating the binding peptide to an agent to  
CC form a binding peptide conjugate and exposing a target to the binding  
CC peptide conjugate, where the binding peptide conjugate binds to the  
CC target. An enzymatic composition comprising the binding peptide is useful  
CC for modifying a tea or wine stain on a fabric or a surface, which  
CC involves contacting the surface having tea or wine stain with the  
CC enzymatic composition. The surface is preferably a ceramic surface, hair  
CC or skin. The binding peptide is useful for producing enzymatic  
CC compositions for use in detergent or cleaning compositions such as for  
CC removing food stains on fabrics or removing food stains on surfaces such  
CC as ceramic and teeth. The binding peptide is useful in the wine ageing  
CC process, where a tannin compound is targeted at the early stage of wine  
CC fermentation process for reducing or eliminating astringency of tannins  
CC in the wine. The binding peptide, combined with a bleaching agent is  
CC useful for delivering a bleaching agent to stained teeth for bleaching.  
CC The binding peptide and binding conjugate are useful in personal care  
CC products such as creams, lotions and ointments. The binding peptide is  
CC useful for modifying astringency in skin. This sequence represents a  
CC binding peptide which specifically binds to tea stains on ceramics for  
CC use in the stain removal process of the invention.  
XX  
SQ Sequence 7 AA;  
XX  
Query Match 8.6%; Score 5; DB 8; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 5 PPSPL 9  
|||  
1 PPSPL 5  
XX  
Db 1 PPSPL 5  
XX  
RESULT 13  
ADP74892  
ID ADP74892 standard; peptide; 7 AA.  
XX  
AC ADP74892;  
XX  
DT 23-SEP-2004 (first entry)  
XX  
DE Parapoxvirus ovis genome ORF-encoding N-terminal peptide #42.  
XX  
XX Parapoxvirus ovis; PPVO; virucide; cytostatic; anti-inflammatory;  
KW immunosuppressive; anti-allergic; gene therapy; viral infection;  
KW non-viral infection; proliferative disease; inflammatory disease;  
KW allergic disease; autoimmune disease; open reading frame; ORF.  
XX  
OS Parapoxvirus.  
XX  
FN WO2004054614-A1.  
XX  
PD 01-JUL-2004.  
XX  
PF 17-DEC-2002; 2002MO-EP014402.  
XX  
PR 17-DEC-2002; 2002MO-EP014402.  
XX  
PA (FARB ) BAYER HEALTHCARE AG.  
XX

PI Weber O, Friederichs SM, Slegling A, Schlapp T, Mercer AA;  
PI Fleming SB;  
XX  
XX WPI; 2004-488011/46.  
XX  
PT New polynucleotide coding for a parapoxvirus ovis (PPVO) genome, useful  
PT for preparing a composition for treating viral or non-viral infections,  
PT or proliferative, inflammatory, allergic or autoimmune diseases.  
XX  
PS Example 4, Page 33; 103pp; English.  
XX  
XX This invention relates to a novel purified and isolated polynucleotide  
CC which comprises a sequence having 137560 bp derived from the Parapoxvirus  
CC ovis (PPVO) and the open reading frames (ORFs) contained within it. The  
CC invention may be useful for the production of compounds with a virucide,  
CC cytostatic, anti-inflammatory, immunosuppressive or anti-allergic activity  
CC whilst the disclosed sequences may be useful for gene therapy. The  
CC invention is useful for preparing a composition for treating viral or non  
CC -viral infections, or proliferative, inflammatory, allergic or autoimmune  
CC diseases. The present sequence is that of a peptide which represents the  
CC N-terminal of a protein encoded by an ORF of the Parapoxvirus genome  
CC sequence of the invention.  
XX  
SQ Sequence 7 AA;  
XX  
Query Match 8.6%; Score 5; DB 8; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 13 LQLLK 17  
|||  
2 LQLLK 6  
XX  
Db 2 LQLLK 6  
XX  
RESULT 14  
AAB06282  
ID AAB06282 standard; peptide; 8 AA.  
XX  
AC AAB06282;  
XX  
DT 03-OCT-2000 (first entry)  
XX  
DE proSKI-1 putative zymogen cleavage site.  
XX  
XX Human, pro-brain-derived neurotrophic factor; proBDNF;  
KW subtilisin-kexin isoenzyme 1; SKI-1; proSKI-1 zymogen processing;  
KW antilipase; cytostatic; vasotropic; SKI-1 inhibitor;  
KW hypercholesterolemia; liver steatosis; Ras-dependent cancer; restenosis;  
KW amyloid protein formation.  
XX  
XX Homo sapiens.  
OS  
OS Mus sp.  
OS Rattus sp.  
XX  
FH Key Location/Qualifiers  
FT Cleavage-site 4..5  
XX  
FN WO200026348-A2.  
XX  
PD 11-MAY-2000.  
XX  
PF 04-NOV-1999; 99WO-CA001058.  
XX  
PR 04-NOV-1998; 98CA-02249648.  
XX  
XX (RECL-) INST RECH CLINIQUES MONTREAL.  
XX  
PI Seidah N, Chretien M, Marcinkiewicz M, Laakonen R, Davignon J;  
XX  
XX WPI; 2000-365601/31.  
XX  
DR Novel soluble proteic fragment of subtilisin-kexin isoenzyme for  
PT producing a polypeptide useful for treating hypercholesterolemia, liver

PT steatosis and amyloidosis, comprises a specific amino acid sequence.  
 XX  
 PS Example 1; Page 25; 119pp; English.  
 CC The present sequence is a possible cleavage site in proSKI-1 which may be  
 CC targeted by a zymogen to form mature subtilisin-kexin isoenzyme 1 (SKI-  
 CC 1). The conserved site was identified by alignment of SKI-1 with other  
 CC subtilases. SKI-1 is a type-1 membrane-bound proteinase which cleaves pro  
 CC -brain-derived neurotrophic factor. Peptides which bind to and are  
 CC cleaved by SKI-1 may be used for monitoring SKI-1 activity, for screening  
 CC inhibitors of SKI-1 activity, or for screening enhancers of SKI-1  
 CC activity. Proteic fragments of SKI-1 which bind to the SKI-1 catalytic  
 CC site may be used as inhibitors of SKI-1 activity. They may be used to  
 CC treat diseases involving overexpression of SKI-1 or SKI-1 substrate. Such  
 CC diseases include hypercholesterolemia, high levels of fatty acids,  
 CC lipids or fatty acid pyrophosphate, liver steatosis, Ras-dependent cancer,  
 CC restenosis and amyloid protein formation  
 XX  
 SQ Sequence 8 AA;  
 Query Match 8.6%; Score 5; DB 3; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 19 SSRLL 23  
 DB 3 SSRLL 7  
 RESULT 15  
 AAR59128  
 ID AAR59128 standard; peptide; 9 AA.  
 XX  
 AC AAR59128;  
 XX  
 DT 16-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 02-MAY-1995 (first entry)  
 XX  
 DE Peptide fragment (1.0882) of HBV binds HLA-A2.1.  
 XX  
 KW antigen; epitope; immunogenic target protein; PSA; HBV; HBVg; EBV; HIV1;  
 KW core antigen; surface antigen; pharmaceutical composition; in vivo;  
 KW ex vivo; therapeutic; diagnostic; MHC class I molecule;  
 KW major histocompatibility complex; HLA-A2.1; 9mer; 10mer; anchor;  
 KW human leukocyte antigen.  
 XX  
 OS Hepatitis B virus; strain adw.  
 XX  
 PN WO9420127-A1.  
 XX  
 PD 15-SEP-1994.  
 XX  
 PF 04-MAR-1994; 94WO-US002353.  
 XX  
 PR 05-MAR-1993; 93US-00027146.  
 PR 04-JUN-1993; 93US-00073205.  
 PR 29-NOV-1993; 93US-00159184.  
 XX  
 PA (CYTE-) CYTEL CORP.  
 XX  
 PI Grey HM, Sette A, Sidney J, Kaet W;  
 XX  
 DR WPI; 1994-302678/37.  
 XX  
 PT Immunogenic peptide(s) having an HLA-A2.1 binding motif - used for  
 PT treatment or prophylaxis of cancer; virus infection or autoimmune  
 PT diseases.  
 XX  
 PS Example 5; Page 101; 138pp; English.  
 XX  
 CC AAR59104-264 are immunogenic 9mer peptides that contain a HLA-A2.1  
 CC binding motif. These peptides bind HLA-A2.1 and have a binding affinity

CC of at least 1% as compared to a reference peptide (AAR71293). AAR59128  
 CC has an IC50 of 0.046 and the sequence occurs at position 1088 in the HBV  
 CC POL protein. The peptides of the invention can induce cytotoxic T  
 CC lymphocytes which can react with target cells. They can be used for the  
 CC treatment or prophylaxis of cancer, eg. prostate cancer or lymphoma, etc.  
 CC (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-OCT-2003 to  
 CC standardise OS field)  
 XX  
 SQ Sequence 9 AA;  
 Query Match 8.6%; Score 5; DB 2; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 34 SIMLL 38  
 DB 5 SIMLL 9

Search completed: October 16, 2005, 05:57:20  
 Job time : 119 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: October 16, 2005, 05:19:43 ; Search time 103.5 Seconds

(without alignments)  
8269.939 Million cell updates/sec

Title: US-10-031-158-13

Perfect score: 1813

Sequence: 1 gggcagagagctgggcacaaa.....attcaaaaatggaagctt 1027

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1859788 seqs, 41671961 residues

Total number of hits satisfying chosen parameters: 3719576

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_n2p.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool\_p/US10031158/runat\_14102005\_135538\_17068/app.query.fasta\_1.1223  
-DB=Published Applications AA -OPMT=faetan -SUFFIX=n2p.rapb -MINMATCH=0.1  
-LOOPCL=0 -LOOEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pcp -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=plc -NORM=ext -HEADSIZE=500 -MINTEN=0  
-MAXLEN=200000000 -USER=US10031158.0CGN\_1\_130.0@runat\_14102005\_135538\_17068  
-NCPU=6 -ICPU=3 -NO\_WMAP -LANG=ENGLISH -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10F\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep:\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1011	55.8	318	10	US-09-957-708-29
2	1005	55.4	318	16	US-10-643-795A-103
3	1005	55.4	318	17	US-10-948-518-103
4	908	50.1	338	17	US-10-847-918-29
5	798	44.0	274	14	US-10-340-536-10
6	558	30.8	105	9	US-09-864-761-47901
7	483.5	26.7	101	9	US-09-925-300-938
8	307	16.9	58	14	US-10-205-623-405
9	307	16.9	58	18	US-10-756-149-5551
10	307	16.9	58	20	US-11-051-454-105
11	196	10.8	534	18	US-10-450-763-35857
12	194	10.7	36	9	US-09-991-548-32
13	138.5	7.6	244	18	US-10-450-763-49648
14	137.5	7.6	216	17	US-10-887-228A-6
15	137.5	7.6	216	17	US-10-887-231-2
16	137	7.6	686	15	US-10-072-012-795
17	136.5	7.5	216	17	US-10-887-231-27
18	134.5	7.4	214	16	US-10-128-520-107
19	134.5	7.4	216	17	US-10-887-231-21
20	133.5	7.4	215	16	US-10-128-520-102
21	133.5	7.4	215	16	US-10-128-520-103
22	133.5	7.4	216	16	US-10-128-520-114
23	133.5	7.4	240	15	US-10-291-265-248
24	133	7.3	218	16	US-10-379-392-159
25	133	7.3	218	16	US-10-379-392-161
26	133	7.3	218	16	US-10-379-392-164
27	133.5	7.3	215	16	US-10-128-520-103
28	133.5	7.3	216	17	US-10-887-231-23
29	131.5	7.3	211	16	US-10-128-520-132
30	131.5	7.3	231	15	US-10-291-265-283
31	131	7.2	229	14	US-10-310-719-36
32	130.5	7.2	211	16	US-10-128-520-99
33	130.5	7.2	213	16	US-10-128-520-138
34	130.5	7.2	214	16	US-10-128-520-122
35	130.5	7.2	214	16	US-10-128-520-124
36	130.5	7.2	214	16	US-10-128-520-127
37	130.5	7.2	214	17	US-10-887-231-17
38	130.5	7.2	215	16	US-10-128-520-135
39	130.5	7.2	216	16	US-10-128-520-120
40	130.5	7.2	216	17	US-10-887-228A-10
41	130.5	7.2	216	17	US-10-887-231-10
42	130.5	7.2	216	17	US-10-887-231-29
43	130.5	7.2	216	17	US-10-887-231-31
44	130.5	7.2	216	17	US-10-887-231-33
45	130.5	7.2	216	17	US-10-887-231-39

## ALIGNMENTS

RESULT 1  
US-09-957-708-29  
Sequence 29, Application US/09957708  
Publication No. US20030031678A1  
GENERAL INFORMATION:  
APPLICANT: Sun, Yongming  
APPLICANT: Recipon, Hervé  
APPLICANT: Cafferey, Robert  
APPLICANT: Ali, Shuifath  
TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific  
FILE REFERENCE: DEX-0239  
CURRENT APPLICATION NUMBER: US/09/957,708  
CURRENT FILING DATE: 2001-09-19  
PRIOR APPLICATION NUMBER: 60/233,746  
PRIOR FILING DATE: 2000-09-19  
NUMBER OF SEQ ID NOS: 40  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 29  
LENGTH: 318





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_n2p model

Run on: October 16, 2005, 04:57:57 ; Search time 37.5 Seconds  
(without alignments)  
4088.778 Million cell updates/sec

Title: US-10-031-158-13  
Perfect score: 1813  
Sequence: 1 ggagcaagagctgggcaaaaa.....atcaaaaaaatgaaagct 1027

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues  
Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_n2p\_model -DEV=xlp  
-Q=/cgn2\_1/USPRO.spool/p/US10031158/runat\_14102005\_135536\_16979/app\_query.fasta\_1.1223  
-DB=Issued Patents AA -QMT=fastran -SUFFIX=n2p.rat -MINMATCH=0.1 -LOOPLC=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdd  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTPRMT=pct -NOR=ext -HEAPSIZE=500 -MUTLEN=0 -MAXLEN=2000000000  
-USER=US10031158 @CGN\_1\_1\_33 @runat\_14102005\_135536\_16979 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUBRY -NEG SCORES=0 -WAIT -DSPBLCK=100 -LONGLOC  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	917	50.6	323	6	5260223-4
2	917	50.6	323	6	5260223-4
3	798	44.0	274	1	US-08-256-964A-19
4	407	22.4	202	6	5189147-6
5	287	15.8	53	4	US-09-513-999C-4756
6	194	10.7	36	3	US-09-048-672A-10
7	137.5	7.6	235	3	US-08-378-939-12
8	129.5	7.0	238	3	US-08-793-450-6
9	127.5	7.0	240	3	US-09-049-672A-11
10	126.5	7.0	236	3	US-08-487-550-10

13	126.5	7.0	236	4	US-09-526-098-10	Sequence 10, Appl
14	126.5	7.0	236	4	US-09-383-916-10	Sequence 10, Appl
15	125.5	6.9	233	3	US-08-523-894-6	Sequence 6, Appl
16	125.5	6.9	233	3	US-09-049-672A-12	Sequence 12, Appl
17	125.5	6.9	235	4	US-09-152-060-70	Sequence 70, Appl
18	125.5	6.9	235	4	US-09-152-060-68	Sequence 88, Appl
19	125.5	6.9	236	3	US-09-049-672A-7	Sequence 7, Appl
20	123.5	6.8	302	4	US-09-789-672A-21	Sequence 21, Appl
21	123.5	6.8	303	4	US-09-509-347-7	Sequence 7, Appl
22	122.5	6.8	234	3	US-08-487-350-2	Sequence 2, Appl
23	122.5	6.8	234	4	US-09-526-098-2	Sequence 2, Appl
24	122.5	6.8	234	4	US-09-383-916-2	Sequence 2, Appl
25	121.5	6.7	293	6	5189147-3	Patent No. 5189147
26	121.5	6.7	293	6	5189147-3	Patent No. 5189147
27	119.5	6.6	215	4	US-09-949-016-10696	Sequence 10696, A
28	118.5	6.5	145	4	US-09-949-016-8908	Sequence 8908, Ap
29	115.5	6.4	306	3	US-09-082-593-2	Sequence 2, Appl
30	114.5	6.3	234	4	US-09-372-425A-4	Sequence 4, Appl
31	114	6.3	236	4	US-09-315-926A-79	Sequence 79, Appl
32	113.5	6.3	200	6	5189147-10	Patent No. 5189147
33	113.5	6.3	200	6	5189147-10	Patent No. 5189147
34	113.5	6.3	234	4	US-09-472-087-17	Sequence 17, Appl
35	113.5	6.3	234	4	US-09-472-087-69	Sequence 69, Appl
36	113.5	6.3	237	1	US-08-398-612A-28	Sequence 28, Appl
37	113.5	6.3	237	1	US-08-398-612A-28	Sequence 28, Appl
38	113.5	6.3	237	2	US-08-491-334A-28	Sequence 28, Appl
39	113.5	6.3	237	3	US-09-027-449-25	Sequence 25, Appl
40	113.5	6.3	237	3	US-08-804-444A-25	Sequence 25, Appl
41	113.5	6.3	237	3	US-09-026-985-25	Sequence 25, Appl
42	113.5	6.3	237	4	US-09-121-952A-25	Sequence 25, Appl
43	113.5	6.3	237	4	US-09-234-340A-25	Sequence 25, Appl
44	113.5	6.3	238	1	US-08-398-613A-28	Sequence 28, Appl
45	113	6.2	224	4	US-09-456-090A-40	Sequence 40, Appl

## ALIGNMENTS

RESULT 1  
5260223-4  
Patent No. 5260223  
APPLICANT: BRENNER, MICHAEL B.; STROMINGER, JACK L.; SEIDMAN,  
JOHN G.; IP, STEPHEN H.; KRANGEL, MICHAEL S.  
TITLE OF INVENTION: METHODS FOR DETECTION OF HUMAN GAMMA  
CELL RECEPTOR  
NUMBER OF SEQUENCES: 4  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/187,698  
FILING DATE: 29-APR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 115,256  
FILING DATE: 29-OCT-1987  
APPLICATION NUMBER: 16,252  
FILING DATE: 19-FEB-1987  
APPLICATION NUMBER: 882,100  
FILING DATE: 03-JUL-1986  
SEQ ID NO: 4  
LENGTH: 323  
5260223-4

Alignment Scores:  
Pred. No.: 2,23e-98  
Score: 917.00  
Percent Similarity: 90.05%  
Best Local Similarity: 86.57%  
Query Match: 50.58%  
DB: 6  
Gaps: 1

US-10-031-158-13 (1-1027) x 5260223-4 (1-323)

QY 25 AAGTATTGGTCCCGAGCAAGCTTATCATTCAGATTAACAAGCTTATGATGAGATGTT 84  
DB 123 LysleupheglyserglyThrleuValValThrleuplyGlnleuAspVal 142

QY 85 TCCCCAGCCCACTATTCTTCTTCAATGCTGTAAGCAAGCTCCAGAGGCTGGA 144  
DB 143 SerProlyspromitrllePheLeuProserllealagutrllyLeuGlnlyalagly 162  
QY 145 ACATACCTTGTCTTCTTGAAGAAATTTTCCCTGATGTTATTAAGTACATTGGAGAA 204  
DB 163 ThrTyreucyLeuLeuGlnlyPhePheProspilietllelyllelletrlglnly 182  
QY 205 AAGAAGCAACACGATTCCTGAGATCCAGAGGGAACACCATGAGACTACAGACACA 264  
DB 183 LysLysSerAntlrilleLeuGlySerGlnGlnlyAntlrmetlyeThrAsnAspThr 202  
QY 265 TACATGAATTTAGCTGTTAACGCTCCAGAAAGTCACTGACCAAGAAACACAGATGT 324  
DB 203 TyrmetySpresertrPleuthrValProGlnGlnSerLeuAspLyglnHlaArgCys 222  
QY 325 ATCGTCAGACATGAGAAATTAATAAAGCGAGTTGATCAAGAAATTAATCTTCTCCAA 384  
DB 223 lleValArgHlsglnAenAenlyAasnGlylleAspGlnGlnlyllePheProProile 242  
QY 385 AAGACG-----GATGTC 396  
DB 243 LysThrAspValThrThrValAspProlyeTyraenTyserLyAspAlaAsnAspVal 262  
QY 397 ATCAATGATGCCCAAGACAAATTTGTCAAAGATGCAATGATACACTGCTGCGAG 456  
DB 263 lleThrmetAspProlyAspAsntrPserLyAspAlaAsnAspThrLeuLeuGln 282  
QY 457 CTCACAAACACCTGATATTACATGATGACCTCTCTGCTCTCAAGAGTGGTCTAT 516  
DB 283 LeuthraenThrserAlaTyTrThrTyreLeuLeuLeuLeuLySerValValTyr 302  
QY 517 TTGGCATCATCACCTGCTGTCTGTAGAAAGACGCTTCTGCTGCAATGAGAGAAA 576  
DB 303 PheAlaIlelleltnrCysCysLeuLeuArgtrAlaPheCysCysAsnGlyGlnly 322  
QY 577 TCA 579  
DB 323 Ser 323

RESULT 2  
5260223-4  
PATENT NO. 5260223  
APPLICANT: BRENNER, MICHAEL B.; STROMINGER, JACK L.; SEIDMAN,  
JOHN G.; IP, STEPHEN H.; KRANGEL, MICHAEL S.  
TITLE OF INVENTION: METHODS FOR DETECTION OF HUMAN GAMMA  
T CELL RECEPTOR  
NUMBER OF SEQUENCES: 4  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/187,698  
FILING DATE: 29-APR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 115,256  
FILING DATE: 29-OCT-1987  
APPLICATION NUMBER: 16,252  
FILING DATE: 19-FEB-1987  
APPLICATION NUMBER: 882,100  
FILING DATE: 03-JUL-1986  
SEQ ID NO: 4  
LENGTH: 323  
5260223-4

Alignment Scores:  
Pred. No.: 2,23e-98 Length: 323  
Score: 917.00 Matches: 174  
Percent Similarity: 90.05% Conservative: 7  
Best Local Similarity: 86.57% Mismatches: 4  
Query Match: 50.58% Indels: 16  
DB: 6 Gaps: 1

US-10-031-158-13 (1-1027) x 5260223-4 (1-323)

QY 25 AAGTATTGGTCCCGGAACAAAGCTTATCATATACGATTAAGCAACTGATGCAATGTT 84

DB 123 LysLeuPheGlySerGlyThrThrLeuValValThrAspLyglnLeuAspAlaAspVal 142  
QY 85 TCCCCAGCCCACTATTCTTCTTCAATGCTGTAAGCAAGCTCCAGAGGCTGGA 144  
DB 143 SerProlyspromitrllePheLeuProserllealagutrllyLeuGlnlyalagly 162  
QY 145 ACATACCTTGTCTTCTTGAAGAAATTTTCCCTGATGTTATTAAGTACATTGGAGAA 204  
DB 163 ThrTyreucyLeuLeuGlnlyPhePheProspilietllelyllelletrlglnly 182  
QY 205 AAGAAGCAACACGATTCCTGAGATCCAGAGGGAACACCATGAGACTACAGACACA 264  
DB 183 LysLysSerAntlrilleLeuGlySerGlnGlnlyAntlrmetlyeThrAsnAspThr 202  
QY 265 TACATGAATTTAGCTGTTAACGCTCCAGAAAGTCACTGACCAAGAAACACAGATGT 324  
DB 203 TyrmetySpresertrPleuthrValProGlnGlnSerLeuAspLyglnHlaArgCys 222  
QY 325 ATCGTCAGACATGAGAAATTAATAAAGCGAGTTGATCAAGAAATTAATCTTCTCCAA 384  
DB 223 lleValArgHlsglnAenAenlyAasnGlylleAspGlnGlnlyllePheProProile 242  
QY 385 AAGACG-----GATGTC 396  
DB 243 LysThrAspValThrThrValAspProlyeTyraenTyserLyAspAlaAsnAspVal 262  
QY 397 ATCAATGATGCCCAAGACAAATTTGTCAAAGATGCAATGATACACTGCTGCGAG 456  
DB 263 lleThrmetAspProlyAspAsntrPserLyAspAlaAsnAspThrLeuLeuGln 282  
QY 457 CTCACAAACACCTGATATTACATGATGACCTCTCTGCTCTCAAGAGTGGTCTAT 516  
DB 283 LeuthraenThrserAlaTyTrThrTyreLeuLeuLeuLeuLySerValValTyr 302  
QY 517 TTGGCATCATCACCTGCTGTCTGTAGAAAGACGCTTCTGCTGCAATGAGAGAAA 576  
DB 303 PheAlaIlelleltnrCysCysLeuLeuArgtrAlaPheCysCysAsnGlyGlnly 322  
QY 577 TCA 579  
DB 323 Ser 323

RESULT 3  
US-08-256-964A-19  
Sequence 19, Application US/08256964A  
PATENT NO. 5723309  
GENERAL INFORMATION:  
APPLICANT: BONNEVILLE, MARC  
TITLE OF INVENTION: PRODUCTION OF SUBUNITS OF SOLUBLE T  
TITLE OF INVENTION: RECEPTORS BY CO-TRANSECTION, USES OF THE PRODUCTS THUS  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Young & Thompson  
STREET: 745 South 23rd Street  
CITY: Arlington  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,964A  
FILING DATE: 14-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 92 14203  
FILING DATE: 25-NOV-1992  
ATTORNEY/AGENT INFORMATION:

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: October 15, 2005, 18:17:13 ; Search time 123.5 Seconds

(without alignments)  
8516.685 Million cell updates/sec

Title: US-10-031-158-13

Perfect score: 1813

Sequence: 1 gggcagaggtgggcagaaaa.....attcaaaaatgaaagtc 1027

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_n2p.model -DEV=xip  
-O=/cgn2.1/USPTO\_epool\_p/US10031158/runat\_14102005\_135536\_16953/app\_query.fasta\_1.1223  
-DB=UniProt\_03 -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPFC=0 -LOOPEXT=0  
-UNITS=bits -START=-1 -END=-1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USERS=US10031158 @CGN 1.1.152 @runat\_14102005\_135536\_16953 -NCPU=6 -ICPU=3  
-DEV -LARGESQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-NO MAP -LARGESQUERY -NEG SCORES=0 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database : UniProt\_03: \*  
1: uniProt\_sprot: \*  
2: uniProt\_trembl: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1003	55.3	315	2	Q6IN96
2	987	54.4	191	2	Q6P5Q5
3	958	52.8	298	2	Q6IN44
4	917	50.6	323	2	Q6P192
5	874	48.2	189	1	TCC_HUMAN
6	654.5	36.1	172	1	TCC2_MOUSE
7	645	35.6	198	2	Q7TN06
8	630	34.7	169	1	TCC3_MOUSE
9	618	34.1	167	1	TCC1_MOUSE
10	593.5	32.7	340	2	Q6PKA9
11	562.5	31.0	190	1	TCC4_MOUSE
12	144.5	8.0	235	2	Q6P192
13	133.5	7.4	230	2	Q6P192
14	130.5	7.2	234	2	Q6P192
15	130.5	7.2	235	2	Q6P192
16	129.5	7.1	235	2	Q6P192

17	129.5	7.1	236	2	Q6NEJ1
18	129.5	7.1	240	2	Q6NFK3
19	128.5	7.1	233	2	Q6NSP4
20	128.5	7.1	233	2	Q6TBC9
21	128.5	7.1	234	2	Q6N355
22	128.5	7.1	234	2	Q6GMV8
23	128.5	7.1	234	2	Q7Z2U7
24	128.5	7.1	236	2	Q6GMX3
25	128.5	7.1	236	2	Q6IP00
26	128.5	7.1	236	2	Q6PIQ7
27	127.5	7.0	208	2	Q6V501
28	126.5	7.0	230	2	Q7Z2U3
29	126.5	7.0	303	2	Q7TND8
30	125.5	6.9	236	2	Q6GMX4
31	124.5	6.9	235	2	Q6PIK1
32	123.5	6.8	233	2	Q6GMW4
33	123.5	6.8	234	2	Q6NS96
34	123.5	6.8	234	2	Q6NS95
35	123.5	6.8	235	2	Q6GMW6
36	123.5	6.8	236	2	Q6EBE1
37	123.5	6.8	236	2	Q6GMV7
38	123.5	6.8	236	2	Q6P5S3
39	120.5	6.6	104	1	LAC1_RAT
40	119.5	6.6	233	2	Q6P169
41	118.5	6.5	235	2	Q6P2J1
42	118.5	6.5	237	2	Q6DHW4
43	117.5	6.5	213	1	TILL_HUMAN
44	117.5	6.5	235	2	Q6P1F2
45	117.5	6.5	294	2	Q6K1Z5

## ALIGNMENTS

### RESULT 1

ID	Q6IN96	PRELIMINARY;	PRT;	315 AA.
AC	Q6IN96;			
DT	05-JUL-2004 (TREMBLrel. 27, Created)			
DT	05-JUL-2004 (TREMBLrel. 27, Last sequence update)			
DT	05-JUL-2004 (TREMBLrel. 27, Last annotation update)			
DE	TRGV9 protein.			
GN	Name=TRGV9;			
OS	Homio sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Blood;			
RX	MEBLIN=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein W.J., Usdin T.B., Toshiyuki S., Carrinci P., Prange C.,			
RA	Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,			
RA	Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hilyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Paley J., Helton E., Keltman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,			
RA	Krzywinski M.I., Skalska U., Smalutz D.B., Schnerch A., Schein J.E.,			
RT	Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Blood;			



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus.n2p model

Run on: October 16, 2005, 04:25:08 ; Search time 37 Seconds

(without alignments)  
5341.331 Million cell updates/sec

Title: US-10-031-158-13

Perfect score: 1813

Sequence: 1 gggcagaagtggtggcagaaaa.....attcaaaaatggaagtc 1027

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Xgapop 6.0, Xgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+n2p.model -DEV=xlp  
-Q=/cgnr/1/USFPO\_pool/p/US10031158/runat\_14102005\_135536\_16965/app.query.fasta\_1.1223  
-DB=FIR\_79 -QFMT=fastan -SUFPT=n2p.rpr -MINMATCH=0.1 -LDOFCL=0 -LDOEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOCMATCH=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10031158 @CGN 1.1 38 @runat\_14102005\_135536\_16965 -NCPU=6 -ICPU=3  
-NO MMAR -LARGOQUERY -NEG SCORES=0 -WAIT -DSFBLD=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELDP=6 -DELEXT=7

Database :  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	917	50.6	323	2	S01895
2	916	50.5	173	2	A26659
3	910	50.2	340	2	S03517
4	874	48.2	189	1	RMWUGC
5	874	48.2	189	2	B26659
6	866	47.8	205	2	B26425
7	814	44.9	304	2	S04663
8	798	44.0	210	2	S33440
9	662.5	36.5	210	2	JC4683
10	654.5	36.1	172	1	RMWSC7
11	642.5	35.4	172	2	D26420
12	630	34.7	169	1	RMWSC2
13	628.5	34.7	171	2	S04831
14	618	34.1	167	1	RMWSC1

15	582.5	32.1	206	2	C49054	T-cell receptor ga
16	562.5	31.0	190	1	RMWSC8	T-cell receptor ga
17	562	31.0	204	2	S36287	T-cell receptor ga
18	553.5	30.5	234	2	B43546	T-cell receptor ga
19	551.5	30.4	246	2	A43546	T-cell receptor ga
20	550	30.3	179	2	S36289	T-cell receptor ga
21	549	30.3	216	2	S36288	T-cell receptor ga
22	270	14.9	52	2	A23240	T-cell receptor ga
23	270	14.9	202	2	S36293	T-cell receptor ga
24	267	14.7	168	2	S36294	T-cell receptor ga
25	263	14.5	210	2	S36297	T-cell receptor ga
26	259	14.3	203	2	S36291	T-cell receptor ga
27	253	14.0	203	2	S23043	T-cell receptor ga
28	241	13.3	185	2	S36290	T-cell receptor ga
29	204	11.3	179	2	S36295	T-cell receptor ga
30	145	8.0	218	2	A36198	T-cell receptor be
31	138.5	7.6	243	2	S25755	Ig lambda chain -
32	137.5	7.6	235	2	S25749	Ig lambda chain -
33	134.5	7.4	208	2	B49444	Ig lambda chain (N
34	133.5	7.4	214	2	PC4156	Ig lambda chain V
35	131.5	7.3	231	2	S25751	Ig lambda chain -
36	130.5	7.2	118	2	A39949	Ig lambda chain J -
37	130.5	7.2	226	2	S25745	Ig lambda chain -
38	129.5	7.1	216	2	S69130	Ig lambda chain (D
39	129.5	7.1	231	2	S25753	Ig lambda chain -
40	129.5	7.1	233	2	S25741	Ig lambda chain -
41	129.5	7.1	240	2	A41797	Ig light chain - s
42	128.5	7.1	118	2	S12442	Ig lambda chain (K
43	128.5	7.1	118	2	S12441	Ig lambda chain (K
44	128.5	7.1	190	2	S25740	Ig lambda chain -
45	128.5	7.1	213	2	S21066	Ig lambda chain V

## ALIGNMENTS

## RESULT 1

S01895  
T-cell receptor gamma chain precursor - human

C1Species: Homo sapiens (man)

C1Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 21-Jan-2000

C1Accession: S01895

R1Hochstenbach, F.; Parker, C.; McLean, J.; Gieselmann, V.; Band, H.; Bank, I.; Chess, L

J. Exp. Med. 168, 761-776, 1988

A1Title: Characterization of a third form of the human T cell receptor gamma/delta.

A1Reference number: S01895; MUID:88316186; PMID:3261778

A1Accession: S01895

A1Molecule type: mRNA

A1Residues: 1-323 <HOC>

A1Cross-references: EMBL:Y00790; NID:G37017; PIDN:CA68744.1; PID:G37018

C1Superfamily: Immunoglobulin V region; Immunoglobulin homology

C1Keywords: glycoprotein; T-cell receptor

F114/Domain: signal sequence #status predicted <SIG>

F115-323/Product: T-cell receptor gamma chain #status predicted <MAT>

F115-120/Region: V segment

F1121-134/Region: J segment

F1135-323/Domain: C region #status predicted <CRE>

F1159-224/Domain: Immunoglobulin homology <IMM>

F1106,200,254,270,276,285/Binding site: carbohydrate (asn) (covalent) #status predicted

## Alignment Scores:

Pred. No.: 9.57e-80  
Score: 917.00  
Percent Similarity: 90.05%  
Best Local Similarity: 86.57%  
Query Match: 50.58%  
Length: 323  
Matches: 174  
Conservative: 7  
Mismatch: 4  
Indels: 16  
Gaps: 1

US-10-031-158-13 (1-1027) x S01895 (1-323)

QY 25 AAGGATTTGGTCCCGAACAAGCTTATCATCAAGTAACAAGCTTGATGAGATGTT 84  
DB 123 Lysleuphegylserglythrleuvalvalthrleuplysghlnleupalaaspval 142

```

QY      85  TCCTCCAGCCGACATTTTCTTCTTCATTTGCTGAAACAAAGCTCCAGAGGCTGCA 144
Db      143  SerProlyserProthrllePheleuProSerllealagIurThryleuGlnlyleagly 162
QY      145  ACATCCTTTGCTCTTGAGAAATTTTCCCTGATGTATTAATCATATTGGAGAA 204
Db      163  ThrTryleuCyseuueuGlnlylePhePheProabPillelyeYalIehIstrpGlnu 182
QY      205  AAGAAAGCAACACGATTTCTGGGATCCAGAGGGGAAACACCATGAGAAGCTAACGACA 264
Db      183  LysIysSerAenThrlleuGlySerGlnGlnGlyAenThrMetIyethrAenapThr 202
QY      265  TACATGAATTTAGCTGTTAACTGCTCCAGAAAAGTCACTGAGCAAAAGACAGATGT 324
Db      203  TyrMetIysPheSerTrpleuThryalProGlnGlnSerIeuaSpIyehIhIargCyS 222
QY      325  ATCGTCAGACATGATTAATAAAGAGGTGATGATCAAGAAATATCTTCTCCATCA 384
Db      223  IleValIargHlsgIuAenAenlyleAenGlyIleAenGlnIuIleIlePheProIle 242
QY      385  AAGAGC-----GATGTC 396
Db      243  LysThrAerValIThrThryalAerProlyserTyraenTyserIyAerPalAaenapVal 262
QY      397  ATCACAATGATCCCAAGACAAATGTTCAAAAGATGCAATGATCACTACTGCTGAG 456
Db      263  IleThrMetAerProlyAerAerAerTrpSerIyAerPalAaenapThrIleuIeunGln 282
QY      457  CTCACAACACCTGCTGATATTAATGATGATACCTCCCTGCTGCTCAAGAGTGCTAT 516
Db      283  LeuThrAenThrSerAlaItyrTrpThryleuIeuleuIeuleuIySerValIyIyr 302
QY      517  TTGGCATATCATCCTGCTGCTGTTAAGAAACAGGCTTTCTGCTGCAATGAGAGAA 576
Db      303  PheAlaIleIleThrCyseIeuleuIyAerGATrAlaPheCyCyAenGlyGlnlySer 322
QY      577  TCA 579
Db      323  Ser 323

RESULT 2
A:26659
T-cell receptor gamma-1 chain C region - human
N/Alternate names: T-cell receptor gamma chain C region (D-P.LL)
C/Species: Homo sapiens (man)
C/Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 21-Jan-2000
C/Accession: A26659; B22340
R/Refname: M.P.; Forster, A.; Rabbits, T.H.
A/Title: Genetic polymorphism and exon changes of the constant regions of the human T-ce
A/Reference number: A94708; MUID:87092294; PMID:2879283
A/Accession: A26659
A/Molecule type: DNA
A/Residues: 1-173 <LEF>
R/Refname: M.P.; Rabbits, T.H.
Nature 316, 464-466, 1985
A/Title: Two tandemly organized human genes encoding the T-cell gamma constant-region se
A/Reference number: A93365; MUID:85268055; PMID:2991773
A/Accession: B22340
A/Molecule type: DNA
A/Residues: 1-52 <LE2>
C/Genetics:
A/Intons: 111/1; 127/1
C/Superfamily: Immunoglobulin C region; Immunoglobulin homology
C/Keywords: T-cell receptor
F:25-90/Domain: Immunoglobulin homology <IM>

Alignment Scores:
Pred. No.: 1.17e-79 Length: 173
Score: 916.00 Matches: 172
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.42% Mismatches: 0
Query Match: 50.52% Indels: 0

```

```

DB:      2      Gaps:      0
US-10-031-158-13 (1-1027) x A26659 (1-173)
QY      61  GATAAACAAGCTGATGAGATGTTTCCCCCAAGCCCAATATTTTCTTCTCAATGCT 120
Db      1  AspIySerGlnIeuaAerPalAerAerAerProlyserProthrllePheleuProSerlleala 20
QY      121  GAAACAAGCTCCAGAAAGCTGGAACATACCTTTGCTCTTGAAGAAATTTTCCCTGAT 180
Db      21  GluThrIySerGlnIyAalaglyThryIeucySeuIeugIuIyPhePheProabP 40
QY      181  GTTATTAAGATACATTGGAGAAAGAAAGACACACAGATTTCTGGGATCCAGAGGG 240
Db      41  ValIeyIstIehIstrpGlnGlnlyleYserAenThrlleuGlySerGlnGlnly 60
QY      241  AACACATGAAGACTCAACGACACATACATGAAATTTAGCTGTTAACGCTCCAGAAAG 300
Db      61  AenThrMetIySerTrpAenAerPthrTrpMetIySerPheSerTrpIeuthrValProGlnly 80
QY      301  TCACCTGACAAAGAACACAGATGATCGTCAGACATGAGAAATTAATAAAGAGTTGAT 360
Db      81  SerIeuaerPlysgIuHlIargCyalIeValIargHlsgIuAenAenlyleAenGlyValAer 100
QY      361  CAAGAAATTTATCTTTCTCCATTAAGACGATGTCATCAATGATGCCAAGACAT 420
Db      101  GlnGlnIleIleIlePheProIleIleThryAerPalIleThrMetAerProlyAerAen 120
QY      421  TGTTCAAAAGATGGAATGATACACTACTGCTGAGAGCTCACAAACACCTGTCATATTAC 480
Db      121  CySerIyAerPalAaenAerPthrIeuleuIeugIuIeunThrAenThrSerAlaItyr 140
QY      481  ATGTACCTCTCTGCTGCTCTCAAGAGTGCTGATTTTGGCATATGACCTGCTGCTG 540
Db      141  MetIyIeuleuIeuleuIeuleuIySerValIyPheAlaIleIleThryCyseIeu 160
QY      541  CTTAGAAAGACGCTTTCTGCTGCAATGAGAGAAATTA 579
Db      161  LeuAerGATrAlaPheCyCyAenGlyGlnlySer 173

RESULT 3
A:503517
T-cell receptor gamma chain precursor (clone pm17c64) - human
C/Species: Homo sapiens (man)
C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jan-2000
C/Accession: S03517
R/Refname: L.; Forster, A.; Clark, D.M.; Boylston, A.W.; Laventr, I.; Rabbits, T.H.
A/Title: Unusual forms of T cell gamma mRNA in a human T cell leukemia cell line: implic
A/Reference number: S03517; MUID:88083067; PMID:2961573
A/Accession: S03517
A/Molecule type: mRNA
A/Residues: 1-340 <TIG>
A/Cross-references: EMBL:X06774; NID:937346; PIDN:CAA29941.1; PID:937347
A/Note: this sequence was determined from the differentiated gene
C/Superfamily: Immunoglobulin V region; Immunoglobulin homology
C/Keywords: T-cell receptor
F:1-10/Domain: signal sequence #status predicted <SIG>
F:21-340/Product: T-cell receptor gamma chain #status predicted <MR>
F:121-116/Domain: V region (V-gamma-8) #status predicted <VR>
F:121-115/Domain: J region (J-gamma-2) #status predicted <JR>
F:136-340/Domain: C region (C-gamma-2) #status predicted <CR>
F:160-225/Domain: Immunoglobulin homology <IM>

Alignment Scores:
Pred. No.: 4.51e-79 Length: 340
Score: 910.00 Matches: 174
Percent Similarity: 83.41% Conservative: 7
Best Local Similarity: 80.18% Mismatches: 4
Query Match: 50.19% Indels: 32
DB:      2      Gaps:      1
US-10-031-158-13 (1-1027) x S03517 (1-340)

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: October 15, 2005, 18:13:26 ; Search time 98 Seconds

(without alignment)  
8106.182 Million cell updates/sec

Title: US-10-031-158-13

Perfect score: 1813

Sequence: 1 gggcagagctgggcaaaaa.....atcaaaaaatgaaagtc 1027

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2105692 segs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p model -DEV=xlp  
-Q=/cgn2\_1/USFto\_spool\_p/US10031158/runat\_14102005\_135535\_16945/app\_query.fasta\_1.1223  
-DB-A=Geneseq\_16Dec04 -QMT=faetan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPLC=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blom62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=plc -NOM=ext -HEAPSIZE=500 -MITLEN=0 -MAXLEN=2000000000  
-USER=US10031158 @CGN\_1\_1\_154 @runat\_14102005\_135535\_16945 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEOUTRY -NEG\_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database : A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1980s:\*  
3: geneseqp2000s:\*  
4: geneseqp2000s:\*  
5: geneseqp2000s:\*  
6: geneseqp2000s:\*  
7: geneseqp2000s:\*  
8: geneseqp2000s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1011	55.8	318	5	AAE24064 Human pro
2	1005	55.4	318	8	ADL06523 Human tum
3	959	52.9	311	7	ADG61825 Human pro
4	917	50.6	323	2	AA323906 TCR gamma
5	916	50.5	221	8	ADR66006 Human pro
6	916	50.5	221	8	ADR66904 Human pro
7	911	50.2	345	8	ADG0876 Human pro
8	908	50.1	338	8	ADP54940 Human pro
9	907	50.0	324	7	ADG07898 Novel pro
10	894	49.3	324	1	AA31885 Deduced s

11	798	44.0	274	2	AA55705 Human lym
12	798	44.0	274	6	ABR82494 Human T c
13	798	44.0	275	7	ADG61829 Human pro
14	658.5	36.3	331	7	ADG61823 Rat Prote
15	658.5	36.3	331	7	ADG61827 Rat Prote
16	649	35.8	304	1	AA31816 Mammalian
17	649	35.8	304	2	AAW01533 Cytoxic
18	649	35.8	304	3	AAW01515 Murine cy
19	595	32.8	111	4	AA666400 Human pro
20	595	30.8	105	4	AB288899 Peptide #
21	558	30.8	105	5	AA67238 Human bon
22	558	30.8	105	5	AB36589 Human pro
23	483.5	26.7	101	3	AA564420 Human TCR
24	307	16.9	58	4	AB663399 Human TCR
25	307	16.9	58	7	ADG75581 Protease
26	307	16.9	58	8	ADP54863 Human PRO
27	293	16.2	67	3	ADP64625 Human onc
28	287	15.8	53	3	AA600675 Human sec
29	242	13.3	57	1	AA31813 Peptide d
30	242	13.3	57	3	AA50116 Murine CT
31	196	10.8	534	4	ABG05498 Novel hum
32	194	10.7	36	2	AA39432 Human T-c
33	194	10.7	36	5	AAU78032 Human T c
34	175	9.7	62	8	ADP64624 Human onc
35	138.5	7.6	244	4	ABG19289 Human lym
36	137.5	7.6	235	4	AA36212 Human lym
37	137	7.6	686	5	AD117259 Human NOV
38	135.5	7.5	240	3	AA396306 Human IGF
39	134.5	7.4	214	6	ABR01470 Human ant
40	133.5	7.4	214	2	AAW27091 Mouse mon
41	133.5	7.4	215	6	ABR01496 Human ant
42	133.5	7.4	215	6	ABR01465 Human ant
43	133.5	7.4	216	6	ABR01477 Human ant
44	133.5	7.4	240	6	AAU14141 Human nov
45	132.5	7.3	215	6	ABR01466 Human ant

#### ALIGNMENTS

RESULT 1	AAE24064	standard; protein; 318 AA.
AC	AAE24064;	
DT	23-SEP-2002	(first entry)
XX		
DE	Human prostate specific protein (PSP) #7.	
XX		
KW	Human; prostate specific protein; PSP; prostate specific nucleic acid;	
KW	Vaccine; transgenic; prostate cancer; gene therapy; transgenic animal;	
KW	PSNA.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200224718-A1.	
XX		
PD	28-MAR-2002.	
XX		
PF	19-SEP-2001; 2001WO-US029386.	
XX		
PR	19-SEP-2000; 2000US-0233746P.	
XX		
PA	(DIAD-) DIADEXUS INC.	
XX		
PI	Sun Y, Recipon H, Caffery R, Ali S;	
XX		
DR	WPI; 2002-471216/50.	
XX		
FT	Novel isolated prostate specific polypeptide useful for identifying,	
PT	diagnosing, monitoring, staging, imaging, and treating prostate cancer	
XX	and non-cancerous disease states in prostate.	

PS Claim 37, Page 200-201; 210pp; English.

CC The invention relates to prostate specific proteins (PSP) and prostate  
CC specific nucleic acids (PSNA). Sequences of the invention are useful for  
CC identifying, diagnosing, monitoring, staging, imaging and treating  
CC prostate cancer and non-cancerous disease states in prostate. They are  
CC also useful for producing engineered prostate tissue for treatment and  
CC research. The PSNA sequences are used in gene therapy and for producing  
CC transgenic animals and cells. The invention is also used as vaccines. The  
CC present sequence is human prostate specific protein of the invention

XX Sequence 318 AA;

SO Alignment Scores:

Pred. No.:	4,48e-102	Length:	318
Score:	1011.00	Matches:	191
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.48%	Mismatches:	0
Query Match:	55.76%	Indels:	0
DB:	5	Gaps:	0

US-10-031-158-13 (1-1027) x AAE24064 (1-318)

```
QY 4 CAAGAGTTGGGCAAAAAATCAAGTATTGGTCCCGAACAAGCTTATCATACAGAT 63
DB 127 GlnGluLeuGlyLysIleValPheGlyProGlyThrLysLeuIleIleThrAsp 146
QY 64 AAACAACTTGATGAGATGTTCCCGCAAGCCCACTATTCTTCTTCAATGCTGAA 123
DB 147 LysGlnLeuAspAlaAspValSerProLysProThrIlePheLeuProSerIleAlaGlu 166
QY 124 ACAAGCTCCAGAAAGCTGGAACATACCTTTGCTCTTCTTGAGAAATTTTCCCTGATGT 183
DB 167 ThrLysLeuGlnLysAlaGlyThrLysLeuGlyLysLeuGlnLysPheProAspVal 186
QY 184 ATTAAGATACATTGGGAAAGAAAGAGCAACAGATTCGGGATCCCAAGAGGGGAAAC 243
DB 187 IleLysIleHisTrpGlnGlnLysLysSerAsnThrIleLeuGlySerGlnGlnLysAsn 206
QY 244 ACCATGAAGACTAACGACATACATACATGAATTTAGCTGTTAAAGGTCGCAAGAAAGTCA 303
DB 207 ThreLysTrnAsnAspThrLysTrpLysPheSerTrpLeuThrValProGlnLysSer 226
QY 304 CTGACAAAGAAACACAGATGATCGTCAGACATGAGAAATAATAACGAGATGATCA 363
DB 227 LeuAspLysGlnLysIleArgCysIleValAlaArgHisGlnAsnAlaLysGlnValAspGln 246
QY 364 GAATTTATCTTCTTCCCAATTAAGCGATGTCATCAATGATGCCAAGACAAATGT 423
DB 247 GluIleIlePheProIleLysThrAspValIleThrMetAspProLysAspAsnCys 266
QY 424 TCAAAAGATGCAATGATACACTACTGCTGAGCTCACAAACACCTCTGATATATACATG 483
DB 267 SerLysAspAlaAsnAspThrLysLeuLeuGlnLeuTrnAsnTrnSerAlaIleTrpMet 286
QY 484 TACCTCTCTGCTCTCTCAAGAGTGTGCTATTGTGCATCATCACTGCTGTGCTT 543
DB 287 TyrLeuLeuLeuLeuLeuLysSerValValTyrPheAlaIleIleThrCysLeuLeu 306
QY 544 AGAAGAACGGCTTTTGTGCTGCAATGAGAGAAATCA 579
DB 307 ArgArgThrAlaPheCysCysAsnGlyGlnLysSer 318
```

RESULT 2

ADL06523 standard; protein; 318 AA.

ADL06523;

20-MAY-2004 (first entry)

XX Human tumour-associated antigenic target (TAT) polypeptide #22.

KW Human; tumour-associated antigenic target; TAT; cell death; tumour;  
KW cancer; cytostatic.

OS Homo sapiens.

XX WO2004016225-A2.

XX 26-FEB-2004.

XX 19-AUG-2003; 2003WO-US025892.

XX 19-AUG-2002; 2002US-0404809P.

XX 21-AUG-2002; 2002US-0405645P.

XX 23-SEP-2002; 2002US-0413192P.

XX 15-OCT-2002; 2002US-0419008P.

XX 15-NOV-2002; 2002US-0426847P.

XX 02-JUL-2003; 2003US-0484959P.

XX (GENTH ) GENEENTECH INC.

XX Desauvage PJ, Frantz G, Hillan KJ, Polakis P, Poleon A, Smith V,  
PI Spencer SD, Wu TD, Zhang Z;

XX WPI; 2004-257144/24.

XX N-PSDB; ADL06443.

XX New antibody that binds to a tumor-associated antigenic target (TAT)  
PT polypeptide, useful for preparing a composition for diagnosing or  
PT treating cancer.

XX Claim 2; SEQ ID NO 103; 319pp; English.

XX The present invention relates to the isolation of human tumour-associated  
CC antigenic target (TAT) polynucleotide and polypeptide sequences. Also  
CC disclosed is an antibody that binds to a TAT polypeptide. The antibody is  
CC a monoclonal antibody, an antibody fragment, a chimeric antibody or a  
CC humanised antibody. It is conjugated to a growth inhibitory agent. It is  
CC produced in bacteria or in CHO cells and induces death of a cell to which  
CC it binds. The antibody is useful for preparing a composition for  
CC diagnosing or treating tumours and cancer. The present sequence  
CC represents a human TAT polypeptide of the invention.

XX Sequence 318 AA;

Pred. No.:	2,06e-101	Length:	318
Score:	1005.00	Matches:	190
Percent Similarity:	99.48%	Conservative:	1
Best Local Similarity:	98.96%	Mismatches:	1
Query Match:	55.43%	Indels:	0
DB:	8	Gaps:	0

US-10-031-158-13 (1-1027) x ADL06523 (1-318)

```
QY 4 CAAGAGTTGGGCAAAAAATCAAGTATTGGTCCCGAACAAGCTTATCATACAGAT 63
DB 127 GlnGluLeuGlyLysIleValPheGlyProGlyThrLysLeuIleIleThrAsp 146
QY 64 AAACAACTTGATGAGATGTTCCCGCAAGCCCACTATTCTTCTTCAATGCTGAA 123
DB 147 LysGlnLeuAspAlaAspValSerProLysProThrIlePheLeuProSerIleAlaGlu 166
QY 124 ACAAGCTCCAGAAAGCTGGAACATACCTTTGCTCTTCTTGAGAAATTTTCCCTGATGT 183
DB 167 ThrLysLeuGlnLysAlaGlyThrLysLeuGlyLysLeuGlnLysPheProAspVal 186
QY 184 ATTAAGATACATTGGGAAAGAAAGAGCAACAGATTCGGGATCCCAAGAGGGGAAAC 243
DB 187 IleLysIleHisTrpGlnGlnLysLysSerAsnThrIleLeuGlySerGlnGlnLysAsn 206
QY 244 ACCATGAAGACTAACGACATACATACATGAATTTAGCTGTTAAAGGTCGCAAGAAAGTCA 303
DB 207 ThreLysTrnAsnAspThrLysTrpLysPheSerTrpLeuThrValProGlnLysSer 226
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: October 15, 2005, 17:17:31 ; Search time 830 Seconds  
(without alignments)  
8673.659 Million cell updates/sec

Title: US-10-031-158-13

Perfect score: 1027

Sequence: 1 gggcagaagtcgggcaaaaa.....attcaaaaatgaagtc 1027

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 8765912 seqs, 3504951483 residues

Total number of hits satisfying chosen parameters: 17531824

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09C\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/2/pubpna/US10I\_NEW\_PUB.seq.\*  
23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/2/pubpna/US11A\_NEW\_PUB.seq.\*  
25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1025.4	99.8	1027	US-10-205-823-404	Sequence 404, App
2	1025.4	99.8	1027	US-10-756-149-2859	Sequence 2859, App
3	1025.4	99.8	1027	US-11-051-454-404	Sequence 404, App
4	1012.2	98.6	1395	US-10-643-795A-23	Sequence 23, App1
5	1012.2	98.6	1395	US-10-948-518-23	Sequence 23, App1
6	966.6	94.1	1799	US-10-101-510-492	Sequence 492, App
7	944	91.9	1418	US-09-957-708-16	Sequence 16, App1

8	906	88.2	1155	9	US-09-925-300-58	Sequence 58, App1
9	900.6	87.7	1421	9	US-09-954-456-317	Sequence 317, App
10	900.6	87.7	1421	21	US-10-843-641A-3344	Sequence 3344, App
11	836.4	81.4	1586	10	US-09-960-706-676	Sequence 676, App
12	836.4	81.4	1586	10	US-09-873-319-424	Sequence 424, App
13	836.4	81.4	1586	21	US-10-847-918-14	Sequence 14, App1
14	795.8	77.5	1162	20	US-10-357-930-21517	Sequence 21517, A
15	795.8	77.5	1162	20	US-10-357-930-24142	Sequence 24142, A
16	795.8	77.5	1162	20	US-10-357-930-24529	Sequence 24529, A
17	795.8	77.5	1162	20	US-10-357-930-24533	Sequence 24533, A
18	795.8	77.5	1162	20	US-10-357-930-24840	Sequence 24840, A
19	795.8	77.5	1162	20	US-10-357-930-25151	Sequence 25151, A
20	795.8	77.5	1162	20	US-10-357-930-27359	Sequence 27359, A
21	795.8	77.5	1162	20	US-10-357-930-27740	Sequence 27740, A
22	619	60.3	924	20	US-10-357-930-25081	Sequence 25081, A
23	589	57.4	124933	22	US-10-737-082-97	Sequence 97, App1
24	589	57.4	124933	22	US-10-765-780-97	Sequence 97, App1
25	497.2	48.4	1316	20	US-10-357-930-25082	Sequence 25082, A
26	497.2	48.4	1316	20	US-10-357-930-26474	Sequence 26474, A
27	471.6	45.9	486	20	US-10-357-930-25002	Sequence 25002, A
28	450.8	43.9	825	16	US-10-340-536-29	Sequence 29, App1
29	446.8	43.5	784	20	US-10-357-930-21887	Sequence 21887, A
30	444	43.2	511	9	US-09-759-143-72	Sequence 72, App1
31	444	43.2	511	9	US-09-780-663-72	Sequence 72, App1
32	444	43.2	511	9	US-09-030-605-72	Sequence 72, App1
33	444	43.2	511	9	US-09-822-827-72	Sequence 72, App1
34	444	43.2	511	9	US-09-115-453-72	Sequence 72, App1
35	444	43.2	511	9	US-09-232-880-72	Sequence 72, App1
36	444	43.2	511	9	US-09-895-793-72	Sequence 72, App1
37	444	43.2	511	9	US-09-895-814-72	Sequence 72, App1
38	444	43.2	511	13	US-10-012-886-72	Sequence 72, App1
39	444	43.2	511	14	US-10-010-940-72	Sequence 72, App1
40	444	43.2	511	16	US-10-144-678A-72	Sequence 72, App1
41	444	43.2	511	16	US-10-294-025-72	Sequence 72, App1
42	444	43.2	511	19	US-10-688-838-72	Sequence 72, App1
43	414.6	40.4	569	20	US-10-357-930-43494	Sequence 43494, A
44	408.8	39.8	539	20	US-10-357-930-22784	Sequence 22784, A
45	408.8	39.8	539	20	US-10-357-930-28636	Sequence 28636, A

#### ALIGNMENTS

RESULT 1  
US-10-205-823-404  
Sequence 404, Application US/10205823  
Publication No. US20030108963A1  
GENERAL INFORMATION:  
APPLICANT: Schlegel, Robert  
APPLICANT: Monahan, John B.  
APPLICANT: Monahan, Wilson O.  
APPLICANT: Ganavatharu, Manjula  
APPLICANT: Gorbacheva, Bella  
APPLICANT: Hoersch, Sebastian  
APPLICANT: Kamatkar, Shubhangt  
APPLICANT: Monsey, Angela M.  
APPLICANT: Glatt, Karen  
APPLICANT: Zhao, Xumei  
APPLICANT: Anderson, Dusein  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
THERAPY OF PROSTATE CANCER  
FILE REFERENCE: MRI-044  
CURRENT APPLICATION NUMBER: US/10/205,823  
CURRENT FILING DATE: 2002-07-25  
PRIOR APPLICATION NUMBER: 60/307,982  
PRIOR FILING DATE: 2001-07-25  
PRIOR APPLICATION NUMBER: 60/314,356  
PRIOR FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/325,020  
PRIOR FILING DATE: 2001-09-25  
PRIOR APPLICATION NUMBER: 60/341,746  
PRIOR FILING DATE: 2001-12-12

PRIOR APPLICATION NUMBER: 60/362,158  
PRIOR FILING DATE: 2002-03-05  
NUMBER OF SEQ ID NOS: 455  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 404  
LENGTH: 1027  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-205-823-404

Query Match 99.8%; Score 1025.4; DB 15; Length 1027;  
Best Local Similarity 99.9%; Pred. No. 4.5e-294;  
Matches 1026; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 GGGCAAGAGTTGGGCAAAAAATCAAGTATTGTCGCCGGAACAAAGCTTATCATTA 60
DB 1 GGGCAAGAGTTGGGCAAAAAATCAAGTATTGTCGCCGGAACAAAGCTTATCATTA 60
QY 61 GATTAACAATTGATGACATGTTTCCCGCAAGCCCACTATTTTCTTCTCAATTGCT 120
DB 61 GATTAACAATTGATGACATGTTTCCCGCAAGCCCACTATTTTCTTCTCAATTGCT 120
QY 121 GAAACAAAGCTCCAGAAAGCTGGAACATACCTTTGCTTCTGAGAAATTTTCCCTGAT 180
DB 121 GAAACAAAGCTCCAGAAAGCTGGAACATACCTTTGCTTCTGAGAAATTTTCCCTGAT 180
QY 181 GTTATTAAAGATCATTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 181 GTTATTAAAGATCATTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 AACACCAAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 241 AACACCAAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 TCACCTGAGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 301 TCACCTGAGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 361 CAAGAAATTTATCTTCTCCCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 CAAGAAATTTATCTTCTCCCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 TGTTCAAAAGATGCAAAATATACATACCTGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 TGTTCAAAAGATGCAAAATATACATACCTGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 ATGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
DB 481 ATGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
QY 541 CTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 541 CTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 601 GGGCAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
DB 601 GGGCAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
QY 661 TTTCTTCTGAGTTGGGCAATTTGAGTCTCATGTTGATCTATCTATCTATTTGTA 720
DB 661 TTTCTTCTGAGTTGGGCAATTTGAGTCTCATGTTGATCTATCTATCTATTTGTA 720
QY 721 TTAACGTTTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 721 TTAACGTTTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 AGCCAGAGAGATCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 781 AGCCAGAGAGATCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 841 AAATAGCGCTCTATATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB 841 AAATAGCGCTCTATATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
```

```
QY 901 TTTCTTAAATCAGATTAAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 901 TTTCTTAAATCAGATTAAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 961 TTGCTAGTTGAATTAATGAGTGTGTTTCCGTATTAAGCAAAATTAATTTAA 1020
DB 961 TTGCTAGTTGAATTAATGAGTGTGTTTCCGTATTAAGCAAAATTAATTTAA 1020
QY 1021 AAAAGTT 1027
DB 1021 AAAAGTT 1027
```

RESULT 2  
US-10-756-149-2859  
Sequence 2859, Application US/10756149  
Publication No. US20050181375A1  
GENERAL INFORMATION:  
APPLICANT: Aiz, Natasha  
TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER  
FILE REFERENCE: file  
CURRENT APPLICATION NUMBER: US/10/756,149  
CURRENT FILING DATE: 2004-01-12  
NUMBER OF SEQ ID NOS: 5818  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 2859  
LENGTH: 1027  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-10-756-149-2859

Query Match 99.8%; Score 1025.4; DB 22; Length 1027;  
Best Local Similarity 99.9%; Pred. No. 4.5e-294;  
Matches 1026; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 GGGCAAGAGTTGGGCAAAAAATCAAGTATTGTCGCCGGAACAAAGCTTATCATTA 60
DB 1 GGGCAAGAGTTGGGCAAAAAATCAAGTATTGTCGCCGGAACAAAGCTTATCATTA 60
QY 61 GATTAACAATTGATGACATGTTTCCCGCAAGCCCACTATTTTCTTCTCAATTGCT 120
DB 61 GATTAACAATTGATGACATGTTTCCCGCAAGCCCACTATTTTCTTCTCAATTGCT 120
QY 121 GAAACAAAGCTCCAGAAAGCTGGAACATACCTTTGCTTCTGAGAAATTTTCCCTGAT 180
DB 121 GAAACAAAGCTCCAGAAAGCTGGAACATACCTTTGCTTCTGAGAAATTTTCCCTGAT 180
QY 181 GTTATTAAAGATCATTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 181 GTTATTAAAGATCATTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 AACACCAAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 241 AACACCAAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 TCACCTGAGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 301 TCACCTGAGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 361 CAAGAAATTTATCTTCTCCCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 CAAGAAATTTATCTTCTCCCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 TGTTCAAAAGATGCAAAATATACATACCTGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 TGTTCAAAAGATGCAAAATATACATACCTGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 ATGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
DB 481 ATGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
```



QY 385 AAGAC-----GGATGTC 396  
 DB 763 AAGACAGATGTCACACAGTGGATCCCAATTAATTATCAAGAGTCAATATATGTC 822  
 QY 397 ATCACAATGATGCCAAGACAAATTGTTCAAAAAGATGCAATGATACACTACTGCTGCAG 456  
 DB 823 ATCACAATGATGCCAAGACAAATTGTTCAAAAAGATGCAATGATACACTACTGCTGCAG 882  
 QY 457 CTCACAAACACCTGTCATATTAATGATGATACCTCTCTGCTGCTCAAGAGTGTCTAT 516  
 DB 883 CTCAGAAAACCTGTCATATTAATGATGATACCTCTCTGCTGCTCAAGAGTGTCTAT 942  
 QY 517 TTGGCATATCACTGCTGTCTGTTAGAAAGACGGCTTTCTGCTGCAATGAGAGAA 576  
 DB 943 TTGGCATATCACTGCTGTCTGTTAGAAAGACGGCTTTCTGCTGCAATGAGAGAA 1002  
 QY 577 TCATACAGACGGTGGCACAAGAGGCCATCTTTCTCTCATCGTTATTTGCTCAGAG 636  
 DB 1003 TCATACAGACGGTGGCACAAGAGGCCATCTTTCTCTCATCGTTATTTGCTCAGAG 1062  
 QY 637 CGTC 640  
 DB 1063 CGTC 1066

## RESULT 2

5260223-3  
 Patent No. 5260223  
 APPLICANT: BRENNER, MICHAEL B.; STROMINGER, JACK L.; SEIDMAN, JOHN G.; JIP, STEPHEN H.; KRANGEL, MICHAEL S.  
 TITLE OF INVENTION: METHODS FOR DETECTION OF HUMAN GAMMA T CELL RECEPTOR  
 NUMBER OF SEQUENCES: 4  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/187,698  
 FILING DATE: 29-APR-1988  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 115,256  
 FILING DATE: 29-OCT-1987  
 APPLICATION NUMBER: 16,252  
 FILING DATE: 19-FEB-1987  
 APPLICATION NUMBER: 882,100  
 FILING DATE: 03-JUL-1986  
 SEQ ID NO:3:  
 LENGTH: 1080  
 5260223-3

Query Match 50.1%; Score 514.8; DB 6; Length 1080;  
 Best Local Similarity 88.7%; Pred. No. 9.9e-154;  
 Matches 589; Conservative 0; Mismatches 27; Indels 48; Gaps 1;

QY 25 AAGTATTGATCCCGAACAAGCTTATCATTAAGATMAACAATTGATGAGATGTT 84  
 DB 403 AAACCTTTGGCAGTGGAAACAACCTGTTGTCACAGATMAACAATTGATGAGATGTT 462  
 QY 85 TCCCCCAAGCCCACTATTTTCTCTCTTAATTGCTGAAACAAGCTCAGAGGCTGGA 144  
 DB 463 TCCCCCAAGCCCACTATTTTCTCTCTTAATTGCTGAAACAAGCTCAGAGGCTGGA 522  
 QY 145 ACATACCTTGTCTCTTGAAGAAATTTTCCCTGATGTTAATTAAGTACATTGGAGAGA 204  
 DB 523 ACATACCTTGTCTCTTGAAGAAATTTTCCCTGATGTTAATTAAGTACATTGGAGAGA 582  
 QY 205 AAGAAAGACACGATTTCTGGGATCCAGAGGGGAAACACCATGAGAGTAAAGACACA 264  
 DB 583 AAGAAAGACACGATTTCTGGGATCCAGAGGGGAAACACCATGAGAGTAAAGACACA 642  
 QY 265 TACATGAATTTTACGCTGTTAAACGGTCCAGAAAAGTCACTGAGACAAAGACACGATGT 324  
 DB 643 TACATGAATTTTACGCTGTTAAACGGTCCAGAAAAGTCACTGAGACAAAGACACGATGT 702  
 QY 325 ATCGTCAGATGAGATTAATAAAGGAGTTGATCAAGAAATTAATCTTCTCCATTA 384

DB 703 ATCGTCAGATGAGATTAATAAAGGAGTTGATCAAGAAATTAATCTTCTCCATTA 762  
 QY 385 AAGAC-----GGATGTC 396  
 DB 763 AAGACAGATGTCACACAGTGGATCCCAATTAATTATCAAGAGTCAATATATGTC 822  
 QY 397 ATCACAATGATGCCAAGACAAATTGTTCAAAAAGATGCAATGATACACTACTGCTGCAG 456  
 DB 823 ATCACAATGATGCCAAGACAAATTGTTCAAAAAGATGCAATGATACACTACTGCTGCAG 882  
 QY 457 CTCACAAACACCTGTCATATTAATGATGATACCTCTCTGCTGCTCAAGAGTGTCTAT 516  
 DB 883 CTCAGAAAACCTGTCATATTAATGATGATACCTCTCTGCTGCTCAAGAGTGTCTAT 942  
 QY 517 TTGGCATATCACTGCTGTCTGTTAGAAAGACGGCTTTCTGCTGCAATGAGAGAA 576  
 DB 943 TTGGCATATCACTGCTGTCTGTTAGAAAGACGGCTTTCTGCTGCAATGAGAGAA 1002  
 QY 577 TCATACAGACGGTGGCACAAGAGGCCATCTTTCTCTCATCGTTATTTGCTCAGAG 636  
 DB 1003 TCATACAGACGGTGGCACAAGAGGCCATCTTTCTCTCATCGTTATTTGCTCAGAG 1062  
 QY 637 CGTC 640  
 DB 1063 CGTC 1066

## RESULT 3

US-08-256-964A-18  
 Sequence 18, Application US/08256964A  
 Patent No. 5723309

GENERAL INFORMATION:  
 APPLICANT: BONNEVILLE, MARC  
 TITLE OF INVENTION: PRODUCTION OF SUBUNITS OF SOLUBLE T  
 TITLE OF INVENTION: RECEPTORS BY CO-TRANSFECTION, USBS OF THE PRODUCTS THUS  
 TITLE OF INVENTION: OBTAINED  
 NUMBER OF SEQUENCES: 19  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Young & Thompson  
 STREET: 745 South 23rd Street  
 CITY: Arlington  
 STATE: VA  
 COUNTRY: USA  
 ZIP: 22202  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/256,964A  
 FILING DATE: 14-SEP-1994  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: FR 92 14203  
 FILING DATE: 25-NOV-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: PATCH, Andrew J.  
 REGISTRATION NUMBER: Reg. No. 5723309 32,925  
 REFERENCE/DOCKET NUMBER: BE 94/449  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 703/521-2297  
 TELEFAX: 703/685-0573  
 TELEX: 248425  
 INFORMATION FOR SEQ ID NO: 18:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 825 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: CDS

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comphen Ltd.

OM nucleic - nucleic search, using bw model

Run on: October 15, 2005, 15:47:50 ; Search time 3341 Seconds  
(without alignments)  
11700.690 Million cell updates/sec

Title: US-10-031-158-13

Perfect score: 1027

Sequence: 1 gggcagaagctggcgcaaaaa.....attcaaaaatgtaaacgt 1027

Scoring table: IDENTITY\_NUC

Gap 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:  
1: gb\_esc1:  
2: gb\_esc2:  
3: gb\_hic:  
4: gb\_esc3:  
5: gb\_esc4:  
6: gb\_esc5:  
7: gb\_esc6:  
8: gb\_gsa1:  
9: gb\_gsa2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	963.4	93.8	3533	3	BC030554 Homo sapi
2	879.2	85.6	1510	3	BC017861 Homo sapi
3	840.8	81.9	1757	3	BC039116 Homo sapi
4	836.2	81.4	914	5	BQ934484 AGENCOURT
5	634.4	61.8	821	2	BF679123 602153390
6	627.8	60.2	997	2	BF678971 602153608
7	618.6	60.2	636	1	AI972955 wt45g05.x
8	613.8	59.8	638	1	AI685999 tE91804.x
9	612.4	59.6	902	2	BF674533 602157811
10	603.6	58.8	959	7	CN645427 ILILMIGEN
11	590.2	57.5	703	4	BG217853 RST37574
12	589.6	57.4	629	1	AI768834 wJ03h04.x
13	584	56.9	700	6	CA449324 UT-H-E11-
14	580.4	56.5	582	2	BE326754 ht64405.x
15	578.8	56.4	613	4	BI905966 603062856
16	571.4	55.6	719	2	BF681238 602155502
17	566.2	55.1	670	2	BF677916 602084736
18	564	54.9	789	5	BUI99143 DCBCTA08
19	559.4	54.5	629	2	BF679165 602153737
20	556.8	54.2	1058	4	BM544213 AGENCOURT
21	554.6	54.0	881	2	BF677648 602085529
22	551	53.7	849	2	BF679284 602153477
23	547.6	53.3	701	9	AG175832 Pan trogl
24	547.2	53.3	803	2	BF680840 602156088

25	542	52.8	857	2	BF674457	BF674457 602137231
26	538.6	52.4	861	2	BF681385	BF681385 602156579
27	532.4	51.8	665	1	AA569813	AA569813 nm46c02.s
28	530.8	51.7	600	1	AI244217	AI244217 qj86c05.s
29	517.6	50.4	1055	4	BF644804	BF644804 602268596
30	491.8	47.9	927	2	BF679476	BF679476 602153280
31	490.4	47.8	503	1	AI253507	AI253507 aq3c11.x
32	490	47.5	525	5	BX112931	BX112931 BK112931
33	489.2	47.6	572	2	BF370026	BF370026 MR3-FN000
34	488.4	47.6	513	1	AI823897	AI823897 wj28e10.x
35	466	45.4	479	2	AM575113	AM575113 UI-HF-BKO
36	448.6	43.7	529	2	AA508306	AA508306 0933h07.s
37	447.8	43.6	596	1	AV715641	AV715641 AV715641
38	435.6	42.4	442	1	AI718234	AI718234 a842h05.x
39	434.8	42.3	497	1	AA654656	AA654656 nt76c02.s
40	430	41.9	870	2	BF678129	BF678129 602085181
41	429.2	41.8	512	1	AA657507	AA657507 nt64g07.s
42	428	41.7	460	2	BE772783	BE772783 RCL-FT013
43	427	41.6	681	1	AV714756	AV714756 AV714756
44	418.8	40.8	865	1	AI557112	AI557112 PT2.1.13
45	417	40.6	523	2	BF675605	BF675605 602083495

## ALIGNMENTS

RESULT 1	BC030554	3533 bp	mRNA	linear	HTC 20-MAY-2002
LOCUS	BC030554				
DEFINITION	Homo sapiens, similar to T cell receptor gamma locus, clone IMAGE:521435, mRNA.				
ACCESSION	BC030554				
VERSION	BC030554.1	GI:20988582			
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	1 (bases 1 to 3533)				
TITLE	Strausberg, R.				
JOURNAL	Direct Submission				
COMMENT	Submitted (07-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				

REMARK  
COMMENT  
NHL-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)

Alkner, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Boulfard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Latic, P., Legaspi, R., Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stanciroop, S., Thomas, P.J., Touchman, J.W., Tsurgoun, C., Vogt, J.L., Walker, M.A., Wecherby, K.D., Wiggins, L., Young, A., Zhang, L.-H., and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNM, at: <http://image.llnl.gov>  
Series: IRAX Plates: 64 Row: k Column: 18  
This clone has the following problem: frame shifted.

FEATURES  
source  
1..3533  
/organism="Homo sapiens"  
/mol\_type="mRNA"

## ORIGIN

```

/db_xref="taxon:9606"
/clone="IMAGE:5213435"
/issue_type="Blood, adult leukocytes"
/clone_lib="NIH MGC_118"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

```

```

Query Match      93.8%; Score 963.4; DB 3; Length 3533;
Best Local Similarity 99.5%; Pred. No. 2.5e-253;
Matches 964; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 59 CAGATTAACCACTTGATGAGATGTTTCCCCCAAGCCCACTATTTTCTTCTCAATTG 118
DB 743 CTGATTAACCACTTGATGAGATGTTTCCCCCAAGCCCACTATTTTCTTCTCAATTG 802
QY 119 CTGAACAAGGCTCCAGAGGCTGGAACATACCTTGTCTCTGAGAAATTTTCCCG 178
DB 803 CTGAACAAGGCTCCAGAGGCTGGAACATACCTTGTCTCTGAGAAATTTTCCCG 862
QY 179 ATGTTATTAAGATACATTTGGAGAGAAAGAGACACAGATTCGGATCCAGAGAG 238
DB 863 ATGTTATTAAGATACATTTGGAGAGAAAGAGACACAGATTCGGATCCAGAGAG 922
QY 239 GGAACACCATGAAAGACTACACACATACATGAAATTTGCTGTTAACGCTGCCAGAA 298
DB 923 GGAACACCATGAAAGACTACATACATGAAATTTGCTGTTAACGCTGCCAGAA 982
QY 299 AGTCACTGACAAAGACACAGATGATGTCAGACATGAGATTAATAAAGAGATTG 358
DB 983 AGTCACTGACAAAGACACAGATGATGTCAGACATGAGATTAATAAAGAGATTG 1042
QY 359 ATCAAGAAATTAATCTTCTCTCAATTAAGACGAGATGTCATCAATGATCCCAAGAC 418
DB 1043 ATCAAGAAATTAATCTTCTCTCAATTAAGACAGATGTCATCAATGATCCCAAGAC 1102
QY 419 ATGTTTCAAAAGATGAAATGATACATCTGCTGAGAGTCAAAACACTCTGCAATT 478
DB 1103 ATGTTTCAAAAGATGAAATGATACATCTGCTGAGAGTCAAAACACTCTGCAATT 1162
QY 479 ACATGATACCTCTCTGCTCTCAAGAGAGTGGTCTATTTTCCATCATCAGCTGTC 538
DB 1163 ACATGATACCTCTCTGCTCTCAAGAGAGTGGTCTATTTTCCATCATCAGCTGTC 1222
QY 539 TGCTTAAGAAAGACGCTTTCTGCTGCAATGAGAGAAATCATPAACAGCGTGACAG 598
DB 1223 TGCTTAAGAAAGACGCTTTCTGCTGCAATGAGAGAAATCATPAACAGCGTGACAG 1282
QY 599 GAGGCGCATTTTCCCATCGGTTATGTCCTCTAAGAGGCTCTTCTGAGAGATCTAGTTG 658
DB 1283 GAGGCGCATTTTCCCATCGGTTATGTCCTCTAAGAGGCTCTTCTGAGAGATCTAGTTG 1342
QY 659 GCTTTCTTCTGAGGTTTGGGCAATTCAGTTCTCATGTGTACTATTCTATCATTAATG 718
DB 1343 GCTTTCTTCTGAGGTTTGGGCAATTCAGTTCTCATGTGTACTATTCTATCATTAATG 1402
QY 719 TATTAACGTTTCAACAGTGGGACACAGAGAAACCTCACTCTGTAATTAACATGAGA 778
DB 1403 TATTAACGTTTCAACAGTGGGACACAGAGAAACCTCACTCTGTAATTAACATGAGA 1462
QY 779 ATAGGACAGGCGATCTCCAGACACAACTCTCCAGATTTTCCAGAGCTCTCCAGCAG 838
DB 1463 ATAGGACAGGCGATCTCCAGACACAACTCTCCAGATTTTCCAGAGCTCTCCAGCAG 1522
QY 839 CCAATATAGCGCTGCTATAGTGTAGACATCTGCGGCTTCTAGCCTTGTCCCTCTAG 898
DB 1523 CCAATATAGCGCTGCTATAGTGTAGACATCTGCGGCTTCTAGCCTTGTCCCTCTAG 1582
QY 899 TGTTCTTTAATCAATACCTGCTGAGACCTTTTCAATTTTCAACGCGCTGAAAGAGTCTT 958
DB 1583 TGTTCTTTAATCAATACCTGCTGAGACCTTTTCAATTTTCAACGCGCTGAAAGAGTCTT 1642
QY 959 CTTTCTGATGTAATTAATGATGTGTGTTTTCCTGTAATTAAGCAAAATTAATTAATAA 1018

```

```

DB 1643 CTTTCTGATGTAATTAATGATGTGTGTTTTCCTGTAATTAAGCAAAATTAATTAATAA 1702
QY 1019 TGAAGAAGTT 1027
DB 1703 TGAAGAAGTT 1711

```

```

RESULT 2
BC017861      1510 bp  mRNA  linear  HTC 06-DEC-2001
LOCUS        IMAGE:4687960, mRNA.
DEFINITION   Homo sapiens, similar to T cell receptor gamma locus, clone
ACCESSION    BC017861
VERSION      BC017861.1 GI:17389678
KEYWORDS     HTC.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE    1 (bases 1 to 1510)
AUTHORS      Strausberg, R.
TITLE        Direct Submision
JOURNAL      Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

```

REMARK
COMMENT
Contact: MGC help desk
Email: gcgabs-remail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcdick@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

```

```

FEATURES
source
1..1510
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4687960"
/issue_type="Lung"
/clone_lib="NIH MGC_77"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"

```

```

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov
Series: IMAGE Plate: 36 Row: K Column: 19
This clone has the following problem: frame shifted.
Location/Qualifiers
1..1510

```

## ORIGIN

```

Query Match      85.6%; Score 879.2; DB 3; Length 1510;
Best Local Similarity 93.5%; Pred. No. 2.9e-230;
Matches 948; Conservative 0; Mismatches 18; Indels 48; Gaps 1;

QY 59 CAGATTAACCACTTGATGAGATGTTTCCCCCAAGCCCACTATTTTCTTCTCAATTG 118
DB 464 CTGATTAACCACTTGATGAGATGTTTCCCCCAAGCCCACTATTTTCTTCTCAATTG 523
QY 119 CTGAACAAGGCTCCAGAGGCTGGAACATACCTTGTCTCTGAGAAATTTTCCCG 178
DB 524 CTGAACAAGGCTCCAGAGGCTGGAACATACCTTGTCTCTGAGAAATTTTCCCG 583
QY 179 ATGTTATTAAGATACATTTGGAGAGAAAGAGACACAGATTCGGATCCAGAGAG 238
DB 584 ATATTATTAAGATACATTTGGAGAGAAAGAGACACAGATTCGGATCCAGAGAG 643
QY 239 GGAACACCATGAAAGACTACACACATACATGAAATTTAGCTGTTAACGCTGCCAGAA 298

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: October 15, 2005, 15:44:35 ; Search time 598 seconds

(without alignments)  
10166.512 Million cell updates/sec

Title: US-10-031-158-13  
Perfect score: 1027

Sequence: 1 ggccagagatgggcaaaaa.....attcaaaaatgaagatc 1027

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: N Geneseq\_16Dec04:\*

1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1025.4	99.8	1027	4	AAFS6401 Human TCR
2	1025.4	99.8	1027	10	ADB75580 Prostate
3	1025.4	99.8	1027	13	ADP54862 Human PRO
4	1012.2	98.6	1395	12	AD106443 Human tum
5	966.6	94.1	1799	6	ABZ35381 Human gen
6	944	91.9	1418	6	ADJ38828 Human PSN
7	913.4	88.9	2658	10	ADBE08798 Novel DNA
8	906	88.2	1155	3	AAFI5623 Human pro
9	904	88.0	1560	10	ADBE06987 Novel cod
10	900.6	87.7	1421	6	AB165007 Lung canc
11	900.6	87.7	1421	13	ADP65880 Human pro
12	900.6	87.7	1421	13	ADP66783 Human pro
13	836.4	81.4	1586	6	ABK64529 Human ben
14	836.4	81.4	1586	13	ADP54939 Human PRO
15	833.6	81.2	2658	10	ADBE08798 Novel DNA
16	820	79.8	1046	12	ADP10449 Reference
17	795.8	77.5	1162	5	ABV27727 Human pro
18	795.8	77.5	1162	5	ABV24153 Human pro
19	795.8	77.5	1162	5	ABV24540 Human pro
20	795.8	77.5	1162	5	ABV25162 Human pro

C 21	795.8	77.5	1162	5	ABV21526 Human pro
C 22	795.8	77.5	1162	5	ABV27345 Human pro
C 23	795.8	77.5	1162	5	ABV24544 Human pro
C 24	795.8	77.5	1162	5	ABV24851 Human pro
C 25	619	60.3	924	5	ABV25092 Human pro
C 26	601	58.5	786	13	ADP55163 Human PRO
C 27	521.2	50.7	1080	2	AAQ37617 TCR gamma
C 28	514.8	50.1	1080	1	AAAN91698 Sequence
C 29	497.2	48.4	1316	5	ABV26485 Human pro
C 30	497.2	48.4	1316	5	ABV25093 Human pro
C 31	471.6	45.9	486	5	ABV25013 Human pro
C 32	450.8	43.9	825	5	AAQ66888 Human Lym
C 33	450.8	43.9	825	9	ACF35988 Human T c
C 34	446.8	43.5	784	5	ABV21896 Human pro
C 35	444	43.2	511	2	AAV58549 Prostate
C 36	444	43.2	511	2	AAV61291 CDNA sequ
C 37	444	43.2	511	3	AAAO6312 Human imm
C 38	444	43.2	511	3	ABR71217 Human pro
C 39	444	43.2	511	4	AAH93428 Human pro
C 40	444	43.2	511	4	AAH63520 Human pro
C 41	444	43.2	511	4	AAH02493 Prostate
C 42	444	43.2	511	4	AAH84742 Human pro
C 43	444	43.2	511	5	ACA59329 Prostate
C 44	444	43.2	511	5	AAH10071 Human pro
C 45	444	43.2	511	6	ABL94892 Human Ia-

#### ALIGNMENTS

##### RESULT 1:

AAFS6401 ID AAFS6401 standard; cDNA; 1027 BP.

XX AAFS6401; 12-APR-2001 (first entry)

XX Human TCRgamma alternate reading frame protein coding sequence.

XX Human; TARP; prostate cancer; breast cancer; immunotherapy;

XX T cell receptor gamma alternate reading frame protein; TCRgamma; ss.

XX Homo sapiens.

XX MO200104309-A1.

XX 18-JAN-2001.

XX 12-JUL-2000; 2000WO-US019039.

XX 13-JUL-1999; 99US-0143560P.

XX 01-OCT-1999; 99US-0157471P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Pastran I, Essand M, Lee B, Vasmatazis G, Wolfgang C;

XX WPI; 2001-081050/09.

XX P-PSDB; AAB66399; AAB66400.

XX Isolated T-cell receptor gamma alternate reading frame protein useful for

XX diagnosing and raising an immune response to prostate cancer and breast

XX cancer.

XX Example 1; Fig 1; 85pp; English.

XX The present invention provides the protein and coding sequences of the

XX human T cell receptor alternate reading frame protein (TRAP). This

XX protein is expressed in prostate and breast cancer cells at higher levels

XX than normal and so can be used in the immunotherapy of these cancers, as

XX well as their detection and prevention





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: October 15, 2005, 15:46:50 ; Search time 4554 Seconds  
(without alignments)  
10927.428 Million cell updates/sec

Title: US-10-031-158-13

Perfect score: 1027

Sequence: 1 gggcagaggtcggcagaaaa.....attcaaaaatgaagtc 1027

Scoring table: IDENTITY NUC

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1027	100.0	1027	6	AX074415 Sequence
2	1025.4	99.8	1027	6	AF151103 Homo sapi
3	1011.6	98.5	1825	9	BC072396 Homo sapi
4	1007.2	98.1	1054	9	BC062761 Homo sapi
5	982.4	95.7	1530	9	BC072387 Homo sapi
6	971.4	94.6	1470	9	BC072387 Homo sapi
7	939.6	91.5	958	9	BC039725 Homo sapi
8	902.6	87.8	1655	9	BC039725 Homo sapi
9	900.6	87.7	1421	6	AX332835 Sequence
10	900.6	87.7	1421	9	M13231 Human T-cell
11	875.8	85.3	1013	9	M27331 Homo sapien
12	863.6	84.1	1402	9	M16804 Human T-cell
13	855.8	83.3	1003	9	M27332 Homo sapien
14	836.4	81.4	1586	9	M27332 Homo sapien
15	820	79.8	1046	9	M17323 Homo sapien
16	820	79.8	1046	11	G31759 Homo sapien
17	795.8	77.5	1162	6	CQ489650 Sequence
18	795.8	77.5	1162	6	CQ492275 Sequence
19	795.8	77.5	1162	6	CQ492662 Sequence

C 20	795.8	77.5	1162	6	CQ492666 Sequence
C 21	795.8	77.5	1162	6	CQ492973 Sequence
C 22	795.8	77.5	1162	6	CQ493284 Sequence
C 23	795.8	77.5	1162	6	CQ495492 Sequence
C 24	795.8	77.5	1162	6	CQ495873 Sequence
C 25	731.8	71.3	757	9	HMWTCGCT
C 26	619	60.3	924	6	CQ493214 Sequence
C 27	601	58.5	786	9	HMWTCGAC
C 28	589	57.4	847	9	HMWTCGCD3
C 29	589	57.4	140691	9	AF159056 Homo sapi
C 30	589	57.4	171816	9	AC006033 Homo sapi
C 31	579.6	56.4	586	9	HMWTCGCO
C 32	576.6	56.1	1041	9	AY190025 Macaca mu
C 33	576.6	56.1	66558	2	AC130306 Homo sapi
C 34	575.6	55.0	1041	9	AY190027 Macaca mu
C 35	521.2	50.7	1080	9	HSCRCR
C 36	514.8	50.1	1080	6	AR364341 Sequence
C 37	497.2	48.4	1316	6	CQ493215 Sequence
C 38	497.2	48.4	1316	6	CQ494607 Sequence
C 39	471.6	45.9	486	6	CQ493135 Sequence
C 40	471.6	45.9	600	9	HMWTCGCI
C 41	450.8	43.9	825	6	I89883 Sequence
C 42	450.8	43.9	825	6	HSTCRGSA
C 43	446.8	43.5	784	6	CQ490020 Sequence
C 44	444	43.2	511	6	BD241985 Compound
C 45	444	43.2	511	6	AR237168 Sequence

#### ALIGNMENTS

RESULT 1  
AX074415  
LOCUS: AX074415 1027 bp DNA linear PAT 06-FEB-2001  
DEFINITION: Sequence 13 from Patent WO0104309.  
ACCESSION: AX074415  
VERSION: AX074415.1 GI:12710553  
KEYWORDS:  
SOURCE: Homo sapiens (human)  
ORGANISM: Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
1  
Pastan, I., Essand, M., Lee, B., Vasmatzis, G. and Wolfgang, C.  
T-cell receptor\_g(g) alternate reading frame protein, (tarp) and  
uses thereof  
Patent: WO 0104309-A 13 18-JAN-2001;  
JOURNAL: UNITED STATES GOVERNMENT (US)  
TITLE: Location/Qualifiers

#### FEATURES

source

#### CDS

1..1027  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
74..250  
/note="unassigned protein product; Coding region for PS-TCR  
gamma 1 polypeptide (TAR)"  
/codon\_start=1  
/protein\_id="CAC28473.1"  
/db\_xref="GI:12710554"  
/translation="MQMPSPPLFFQLQKSSRLRHTFVLRNFSMLRYIGKK  
RRATRFMPDRGRTP"  
247..582  
/note="unassigned protein product; Coding region for PS-TCR  
gamma 2 polypeptide (deduced amino acid sequence not  
displayed along with DNA sequence, due to overlapping  
CDS's)"  
/codon\_start=1  
/protein\_id="CAC28474.1"  
/db\_xref="GI:12710555"  
/translation="MKTNIDYMKSKSLTPERKSLDKERCTVRRHNNKNGVDRIIF  
PIKTDVITWDFDNDKSDANDTLILQLTNTSAIYVYLLILKSVYFAITTCCLART  
AFCCNGKRS"

#### ORIGIN

Query Match 100.0%; Score 1027; DB 6; Length 1027;  
 Best Local Similarity 100.0%; Pred. No. 5,8e-261;  
 Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 GGGCAAGAGTTGGGCAAAAAATCAAGTATTGGTCCCGCAACAAAGCTTATCTTACA 60
DB      1 GGGCAAGAGTTGGGCAAAAAATCAAGTATTGGTCCCGCAACAAAGCTTATCTTACA 60
QY      61 GATAAACAACCTGATGACAGATGTTTCCCGCAAGCCACATATTTCTTCCCTCAATTGCT 120
DB      61 GATAAACAACCTGATGACAGATGTTTCCCGCAAGCCACATATTTCTTCCCTCAATTGCT 120
QY      61 GATTAACAACCTGATGACAGATGTTTCCCGCAAGCCACATATTTCTTCCCTCAATTGCT 120
DB      61 GATTAACAACCTGATGACAGATGTTTCCCGCAAGCCACATATTTCTTCCCTCAATTGCT 120
QY      121 GAAACAAAGCTTCAGAGAGCTTGAACATACCTTTGTCTTGAAGAAATTTTCCCTGAT 180
DB      121 GAAACAAAGCTTCAGAGAGCTTGAACATACCTTTGTCTTGAAGAAATTTTCCCTGAT 180
QY      121 GAAACAAAGCTTCAGAGAGCTTGAACATACCTTTGTCTTGAAGAAATTTTCCCTGAT 180
DB      121 GAAACAAAGCTTCAGAGAGCTTGAACATACCTTTGTCTTGAAGAAATTTTCCCTGAT 180
QY      181 GTTATTAAGATACATTTGGGAGAGAGAGACACAGATTTCTGGGATCCCGAGAGGCG 240
DB      181 GTTATTAAGATACATTTGGGAGAGAGAGACACAGATTTCTGGGATCCCGAGAGGCG 240
QY      241 AACACCATGAGAGACTTACGACATACATGAAATTTAGTGGTTAAGCGTCCAGAGAAAG 300
DB      241 AACACCATGAGAGACTTACGACATACATGAAATTTAGTGGTTAAGCGTCCAGAGAAAG 300
QY      301 TCACTGACAAAGAACACAGATGTTTCGTACAGACATGAGAAATATATTAACCGAGTTGAT 360
DB      301 TCACTGACAAAGAACACAGATGTTTCGTACAGACATGAGAAATATATTAACCGAGTTGAT 360
QY      301 TCACTGACAAAGAACACAGATGTTTCGTACAGACATGAGAAATATATTAACCGAGTTGAT 360
DB      301 TCACTGACAAAGAACACAGATGTTTCGTACAGACATGAGAAATATATTAACCGAGTTGAT 360
QY      361 CAAGAAATATATCTTCCCTCCCAATTAAGAGAGATGTCATACAGATGATCCCAACACAT 420
DB      361 CAAGAAATATATCTTCCCTCCCAATTAAGAGAGATGTCATACAGATGATCCCAACACAT 420
QY      421 TGTTCAAAAGATGCAAAATGATACACTACTGCTGACAGCTCACAAACACCTGTCATATAC 480
DB      421 TGTTCAAAAGATGCAAAATGATACACTACTGCTGACAGCTCACAAACACCTGTCATATAC 480
QY      481 ATGTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
DB      481 ATGTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
QY      541 CTGTAAGAAAGCGCTTTCTCTGCGAATGAGAGAAATCATTAACACGCTGGCAACAGGA 600
DB      541 CTGTAAGAAAGCGCTTTCTCTGCGAATGAGAGAAATCATTAACACGCTGGCAACAGGA 600
QY      541 CTGTAAGAAAGCGCTTTCTCTGCGAATGAGAGAAATCATTAACACGCTGGCAACAGGA 600
DB      541 CTGTAAGAAAGCGCTTTCTCTGCGAATGAGAGAAATCATTAACACGCTGGCAACAGGA 600
QY      601 GGGCAATCTTTCTCTCATGCTTATGTCCTTGAAGCGCTCTCTGAGGATCTAGTTGGGC 660
DB      601 GGGCAATCTTTCTCTCATGCTTATGTCCTTGAAGCGCTCTCTGAGGATCTAGTTGGGC 660
QY      601 GGGCAATCTTTCTCTCATGCTTATGTCCTTGAAGCGCTCTCTGAGGATCTAGTTGGGC 660
DB      601 GGGCAATCTTTCTCTCATGCTTATGTCCTTGAAGCGCTCTCTGAGGATCTAGTTGGGC 660
QY      661 TTTCTTTCTGGGTTGGGCAATTCAGTTCTCATGTTGTGATCTATTTATTTATTTGTA 720
DB      661 TTTCTTTCTGGGTTGGGCAATTCAGTTCTCATGTTGTGATCTATTTATTTATTTGTA 720
QY      721 TAAAGGTTTTCACACGAGGCGCACACAGAGAACTCACTCTGTAATTAACATGAGGAT 780
DB      721 TAAAGGTTTTCACACGAGGCGCACACAGAGAACTCACTCTGTAATTAACATGAGGAT 780
QY      781 AGCCAGCGGAGTCTCAGACCAATCTCTCATGTTTTCACAGCTCTCTCCAGCAACCC 840
DB      781 AGCCAGCGGAGTCTCAGACCAATCTCTCATGTTTTCACAGCTCTCTCCAGCAACCC 840
QY      841 AGCCAGCGGAGTCTCAGACCAATCTCTCATGTTTTCACAGCTCTCTCCAGCAACCC 840
DB      841 AGCCAGCGGAGTCTCAGACCAATCTCTCATGTTTTCACAGCTCTCTCCAGCAACCC 840
QY      841 AAATAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
DB      841 AAATAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
QY      901 TTTCTTAATCAGATACATGCTCTGGAAGCTTTCATTTTACAGCCCTTGAAGAGCTTCT 960
DB      901 TTTCTTAATCAGATACATGCTCTGGAAGCTTTCATTTTACAGCCCTTGAAGAGCTTCT 960
QY      961 TTGCTAGTTGAATATGAGTGTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
DB      961 TTGCTAGTTGAATATGAGTGTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020

```

QY 1021 AAAAGTT 1027  
 DB 1021 AAAAGTT 1027

## RESULT 2

AF151103

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_p2n model

Run on: October 16, 2005, 06:48:32 ; Search time 707 Seconds

(without alignments)  
575.070 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 307

Sequence: 1 MOMPSPPLFFPLQLKOS.....RYIGKKRRATRPWDRGRGP 58

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Zgapop 6.0, Zgapext 7.0  
Delep 6.0, Delext 7.0

Searched: 8765912 seqs, 3504951483 residues

Total number of hits satisfying chosen parameters: 17531824

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame.p2n.model -DEV=rlp  
-O=/cg2\_1/USPTO\_spool/p/US10031158/rnat\_14102005\_135741\_18954/app\_query.fasta\_1.199  
-DB=Published Applications NA -QW=faeap -SUFFIX=p2n.tmp -MINMATCH=0.1  
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=spc -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10031158 @CGN 1.1 723 @rnat\_14102005\_135741\_18954  
-NCPUG=6 -ICPU=3 -NO MAP -LARGEJOINT -NEG\_SCORES=-1 -WAIT -DSPELLOC=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-CGAPOP=6 -CGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cg2\_6/ptodata/2/pubna/US07\_PUBCOMB.seq:\*  
2: /cg2\_6/ptodata/2/pubna/PTC\_NEW\_PUB.seq:\*  
3: /cg2\_6/ptodata/2/pubna/US06\_NEW\_PUB.seq:\*  
4: /cg2\_6/ptodata/2/pubna/US06\_PUBCOMB.seq:\*  
5: /cg2\_6/ptodata/2/pubna/US07\_NEW\_PUB.seq:\*  
6: /cg2\_6/ptodata/2/pubna/PTC\_PUBCOMB.seq:\*  
7: /cg2\_6/ptodata/2/pubna/US08\_NEW\_PUB.seq:\*  
8: /cg2\_6/ptodata/2/pubna/US08\_PUBCOMB.seq:\*  
9: /cg2\_6/ptodata/2/pubna/US09\_PUBCOMB.seq:\*  
10: /cg2\_6/ptodata/2/pubna/US09\_PUBCOMB.seq:\*  
11: /cg2\_6/ptodata/2/pubna/US09C\_PUBCOMB.seq:\*  
12: /cg2\_6/ptodata/2/pubna/US09C\_NEW\_PUB.seq:\*  
13: /cg2\_6/ptodata/2/pubna/US10\_PUBCOMB.seq:\*  
14: /cg2\_6/ptodata/2/pubna/US10\_PUBCOMB.seq:\*  
15: /cg2\_6/ptodata/2/pubna/US10C\_PUBCOMB.seq:\*  
16: /cg2\_6/ptodata/2/pubna/US10C\_PUBCOMB.seq:\*  
17: /cg2\_6/ptodata/2/pubna/US10E\_PUBCOMB.seq:\*  
18: /cg2\_6/ptodata/2/pubna/US10E\_PUBCOMB.seq:\*  
19: /cg2\_6/ptodata/2/pubna/US10G\_PUBCOMB.seq:\*  
20: /cg2\_6/ptodata/2/pubna/US10H\_PUBCOMB.seq:\*  
21: /cg2\_6/ptodata/2/pubna/US10I\_PUBCOMB.seq:\*  
22: /cg2\_6/ptodata/2/pubna/US10J\_PUBCOMB.seq:\*  
23: /cg2\_6/ptodata/2/pubna/US11\_PUBCOMB.seq:\*  
24: /cg2\_6/ptodata/2/pubna/US11\_NEW\_PUB.seq:\*  
25: /cg2\_6/ptodata/2/pubna/US11\_NEW\_PUB.seq:\*  
26: /cg2\_6/ptodata/2/pubna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	307	100.0	477	9	US-09-864-761-15428
2	307	100.0	470	9	US-09-864-761-15428
3	307	100.0	539	20	US-10-357-930-22784
4	307	100.0	539	20	US-10-357-930-22784
5	307	100.0	825	16	US-10-340-536-29
6	307	100.0	1027	15	US-10-205-823-404
7	307	100.0	1027	22	US-10-756-159-2859
8	307	100.0	1027	24	US-11-051-454-404
9	307	100.0	1316	20	US-10-357-930-25082
10	307	100.0	1316	20	US-10-357-930-25082
11	307	100.0	1395	20	US-10-643-795A-23
12	307	100.0	1395	21	US-10-948-518-23
13	307	100.0	1418	10	US-09-957-708-16
14	307	100.0	124933	22	US-10-737-082-97
15	307	100.0	124933	22	US-10-765-780-97
16	304	99.0	486	20	US-10-357-930-25002
17	302	98.4	316	9	US-09-864-761-31943
18	298	97.1	569	20	US-10-357-930-43494
19	294	95.8	1421	9	US-09-954-456-317
20	294	95.8	1421	21	US-10-843-641A-3344
21	294	95.8	1799	15	US-10-101-510-492
22	284	92.5	1586	10	US-09-960-706-676
23	284	92.5	1586	10	US-09-873-319-424
24	284	92.5	1586	21	US-10-847-918-14
25	259	84.4	533	20	US-10-357-930-37875
26	245	78.8	1155	9	US-09-925-300-58
27	240	78.2	1022	20	US-10-357-930-23803
28	240	78.2	1022	20	US-10-357-930-23975
29	240	78.2	1022	20	US-10-357-930-23980
30	240	78.2	1022	20	US-10-357-930-24023
31	240	78.2	1022	20	US-10-357-930-24061
32	240	78.2	1022	20	US-10-357-930-24062
33	240	78.2	1022	20	US-10-357-930-24097
34	240	78.2	1022	20	US-10-357-930-24100
35	240	78.2	1022	20	US-10-357-930-24113
36	240	78.2	1022	20	US-10-357-930-25188
37	240	78.2	1022	20	US-10-357-930-25812
38	240	78.2	1022	20	US-10-357-930-29708
39	240	78.2	1022	20	US-10-357-930-29892
40	240	78.2	1022	20	US-10-357-930-29936
41	240	78.2	1022	20	US-10-357-930-29971
42	240	78.2	1022	20	US-10-357-930-29972
43	240	78.2	1022	20	US-10-357-930-30007
44	227	72.9	475	20	US-10-357-930-39216
45	216	70.4	401	20	US-10-357-930-13495

## ALIGNMENTS

RESULT 1  
US-09-864-761-15428  
Sequence 15428, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecmics-X-1  
CURRENT FILING DATE: 2001-05-23  
CURRENT APPLICATION NUMBER: US/09/864,761  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04

```

; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15428
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF159056.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
US-09-864-761-15428

Alignment Scores:
Pred. No.: 9,576-36 Length: 470
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x US-09-864-761-15428 (1-470)
Qy 1 MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLysGlnSerSer 20
Db 203 ATGCAGATGTTCCCAAGCCCACTATTTCTTCCTTCATGCTGAAACAAAGCTCC 262
Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArganPheSerLeuMetLeuLysArgTyr 40
Db 263 AGAAGCGCTGGAACATACCTTGTCTTCTTGAGAAAATTTTCCCTGATGTTATTAAGATAC 322
Qy 41 IleGlyLysLysArgArgAlaThrArgPheTTPAAPPProArgArgGlyThrPro 58
Db 323 ATTGGCAAGAAAAGAAAGACCAACGATTTCTGATCCAGAGGGAACACCA 376

RESULT 2
US-09-864-761-15438
; Sequence 15438, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT FILING DATE: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15438
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF159056.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.3
US-09-864-761-15438

Alignment Scores:
Pred. No.: 9,766-36 Length: 477
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x US-09-864-761-15438 (1-477)
Qy 1 MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLysGlnSerSer 20
Db 212 ATGCAGATGTTCCCAAGCCCACTATTTCTTCCTTCATGCTGAAACAAAGCTCC 271
Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArganPheSerLeuMetLeuLysArgTyr 40
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 16, 2005, 06:07:46 ; Search time 311 Seconds

(without alignments)  
709.652 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 307  
Sequence: 1 MGFPPSPFFFLQLIKOSS.....RYIKRRARTRFMDPRGTP 58

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Command line parameters:

-MODEL=frame+.p2n.model -DRV=x1p  
-Q=/cgr2\_1/USF70.spool\_p/US10031158/runat\_14102005\_135739\_18853/app\_query.faeta\_1.199  
-DB=EST -QFMT=faeta1 -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=us10031158 @CCN 1.1 5180 @runat\_14102005\_135739\_18853 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEOUTERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_est3: \*  
4: gb\_est4: \*  
5: gb\_est5: \*  
6: gb\_est6: \*  
7: gb\_est7: \*  
8: gb\_est8: \*  
9: gb\_est9: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	307	100.0	427	2	BF679325 602153825
2	307	100.0	510	1	AV716204 602153825
3	307	100.0	596	1	AV715641 602153825
4	307	100.0	603	1	AV714189 602153825
5	307	100.0	803	1	AV756492 602153825
6	307	100.0	818	4	BI771070 603059635
7	307	100.0	821	2	BF679123 602153390
8	307	100.0	861	2	BF681385 602156579
9	307	100.0	865	1	A1557112 602156579

10	307	100.0	881	2	BF677648 602085529
11	307	100.0	902	2	BF674593 602137811
12	307	100.0	927	2	BF679476 602153280
13	307	100.0	1058	4	BM544213 AGENCOURT
14	307	100.0	3533	3	BC030554 Homo sapi
15	304	99.0	789	5	BU199143 DCBCL108
16	294	95.8	653	2	BE974171 601680522
17	294	95.8	1005	5	BM919380 AGENCOURT
18	294	95.8	1757	3	BC039116 Homo sapi
19	290	94.5	434	2	BF747659 QV2-BP061
20	290	94.5	497	2	AM087486 X067C01.X
21	290	94.5	681	1	AV714756 AV714756
22	290	94.5	1510	1	BC017861 Homo sapi
23	289	94.1	1055	4	BF694904 602268596
24	288	93.8	523	2	BF675605 602083495
25	284	92.5	362	1	A1758378 tY68C12.X
26	278	90.6	719	2	BF681238 602155502
27	270	87.9	854	2	BF681029 602155433
28	268	87.3	629	2	BF679165 602153737
29	267	87.0	803	2	BF680840 602156088
30	262	85.3	772	4	BI836020 603084004
31	259	84.4	808	2	BF673289 602136139
32	253	82.4	857	2	BF674457 602137231
33	251	81.8	849	2	BF679284 602153477
34	248	80.8	995	7	CO647648 ILLUMIGEN
35	245	79.8	663	6	CB553506 MMSPO044
36	227	73.9	870	2	BF678129 602085181
37	226	73.6	637	6	CB551454 MMSPO013
38	224	73.0	670	2	BF677916 602084736
39	221	72.0	827	4	BG535408 602563069
40	209	68.1	572	2	BF370026 MRS-FN000
41	204.5	66.6	864	4	BI834993 603088885
42	203	65.1	552	5	BP164784 BP164784
43	200	65.1	572	5	BP162849 BP162849
44	200	65.1	888	5	BP162346 BP162346
45	194	63.2	744	7	CR791490 DXF2p468F

#### ALIGNMENTS

RESULT 1  
BF679325 427 bp mRNA EST 21-DEC-2000  
602153825F1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4294566 5',  
DEFINITION  
mRNA sequence.

ACCESSION  
BF679325  
VERSION  
BF679325.1 GI:11953220

KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
1 (bases 1 to 427)  
NIH-MGC http://mgs.nci.nih.gov/.

AUTHORS  
National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE  
Unpublished (1999)

COMMENT  
Contract: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov

Tissue Procurement: CLONTECH Laboratories, Inc.  
cDNA Library Preparation: CLONTECH Laboratories, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

Plate: L1CM1144 row: b column: 07  
High quality sequence stop: 427.

Location/Qualifiers

FEATURES  
source  
1..427  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4294566"

/lab host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH MGC 83"  
/note="Organ: prostate; Vector: pONR-LIB (Clontech); Site\_1: SfilI (ggcgcttcgagc); Site\_2: SfilI (ggccattatgc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGACGGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Alignment Scores:

Pred. No.:	1,22e-30	Length:	427
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
Df:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF679325 (1-427)

Oy 1 MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLeuLysGlnSerSer 20  
Dd 66 ATGCAGATGTTCCCCCAAGCCCATATTTTCCTTCATATGCTAACAAGAAGCTCC 125  
Oy 21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTrp 40  
Dd 126 AGAAGGCTGGAAACATACCCTTGCTCTTGAGAAAATTTTCCCTGATGTATTAAGATAC 185  
Oy 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
Dd 186 ATTGGCAAGAAAAAGAGAACACAGATTCGGAGATCCAGAGGCGGAACACCA 239

RESULT 2

LOCUS	AV716204	510 bp	mRNA	linear	EST 11-OCT-2000
DEFINITION <td>AV716204 DCB Homo sapiens cDNA clone DCBAMG08 5', mRNA sequence.</td> <td></td> <td></td> <td></td> <td></td>	AV716204 DCB Homo sapiens cDNA clone DCBAMG08 5', mRNA sequence.				
VERSION <td>AV716204.1</td> <td>GI:10797721</td> <td></td> <td></td> <td></td>	AV716204.1	GI:10797721			
KEYWORDS <td>EST.</td> <td></td> <td></td> <td></td> <td></td>	EST.				
SOURCE <td>Homo sapiens (human)</td> <td></td> <td></td> <td></td> <td></td>	Homo sapiens (human)				
ORGANISM <td>Homo sapiens</td> <td></td> <td></td> <td></td> <td></td>	Homo sapiens				
REFERENCE <td>Bukacinska; Metzazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 510) Xu X., Gu J., Liu F., Qu J., Zhao M., Li Y., Huang Q., Zhou J., Song H., Gu Y., Yang Y., Guo G., Xiao H., Li N., Qian B., Guo X., Cheng Z., Xu S., Gu W., Tu Y., Jia J., Fu G., Ren S., Zhong M., Lu G., Cheng Z. and Han Z. Homo sapiens cDNA DCB clones Unpublished (2000) Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919 (ex. 45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai.</td> <td></td> <td></td> <td></td> <td></td>	Bukacinska; Metzazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 510) Xu X., Gu J., Liu F., Qu J., Zhao M., Li Y., Huang Q., Zhou J., Song H., Gu Y., Yang Y., Guo G., Xiao H., Li N., Qian B., Guo X., Cheng Z., Xu S., Gu W., Tu Y., Jia J., Fu G., Ren S., Zhong M., Lu G., Cheng Z. and Han Z. Homo sapiens cDNA DCB clones Unpublished (2000) Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919 (ex. 45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai.				
TITLE <td>JOURNAL</td> <td></td> <td></td> <td></td> <td></td>	JOURNAL				
COMMENT <td></td> <td></td> <td></td> <td></td> <td></td>					

FEATURES

source

1..510

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="caxon:9606"  
/clone="DCBAMG08"  
/cell\_type="endometrial cells"  
/dev\_stage="mature"  
/lab\_host="BM25.8"  
/clone\_lib="DCB"

ORIGIN	/notes="Vector: pTribEx2; Site_1: sflIA; Site_2: sflIB"			
Alignment Scores:	1.51e-30	Length:	510	
Pred. No.:	Score:	Matches:	58	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	100.00%	Indels:	0	
DB:	1	Gaps:	0	
US-10-031-158-14 (1-58) x AV715641 (1-510)				
Oy	1	MetGlnMetPheProPheSerProLeuPhePhePheLeuGlnLeuLeuYgInSerSer	20	
Db	159	ATGAGAGTGTTCCTCCCAAGCCCACTATTTTCCTTCCTCAATGCTGAACAAGCTCC	218	
Oy	21	ATGATGLeuGluHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr	40	
Db	219	AGAAAGCTGGAACACTACCTTGCTCTTGGAATAATTTTCCCGAATGTTAAGATAC	278	
Oy	41	IlleGlyLysIstAspArgAlaThrThrArgPheThrPheProArgArgGlyThrPro	58	
Db	279	ATTGGCAAGAAAGACAGACACGATTCCTGGATCCAGAGCGGAACCA	332	
RESULT 3				
LOCUS	AV715641	596 bp	mRNA	linear
DEFINITION	AV715641 DCB Homo sapiens cDNA clone DCBBSB05 5', mRNA sequence.			
ACCESSION	AV715641			
VERSION	AV715641.1	GI:10797158		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
AUTHORS	1 (bases 1 to 596) Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J., Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X., Cheng,Z., Xu,S., Gu,W., Ju,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Cheng,Z. and Han,Z. Homo sapiens cDNA DCB clones Unpublished (2000) Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919 (ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai.			
TITLE	JOURNAL			
COMMENT				
FEATURES				
source	Location/Qualifiers 1..596 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="DCBBSB05" /cell_type="dendritic cells" /dev_stage="mature" /lab_host="Bm25.8" /clone_id="DCB" /note="Vector: pTribEx2; Site_1: sflIA; Site_2: sflIB"			
ORIGIN				
Alignment Scores:	1.82e-30	Length:	596	
Pred. No.:	Score:	Matches:	58	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	100.00%	Indels:	0	
DB:	1	Gaps:	0	
US-10-031-158-14 (1-58) x AV715641 (1-596)				

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 16, 2005, 06:00:36 ; Search time 432 Seconds  
(without alignments)  
794.780 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 307

Sequence: 1 MGFPPSPFLFFFLQLKQSS.....RYGKKRRATRFWDPRGTP 58

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
MODL=frame+ p2n.model -DEV=xlp  
O=/sgn2.1/USPTO\_spool\_p/US10031158/runat\_14102005\_135738\_18834/app\_query.fasta\_1.199  
-DB=N\_Geneseq\_16Dec04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MOD=LOCAL -OUTFMT=pco -NORF=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10031158 @CGN 1.1.708 @runat\_14102005\_135738\_18834 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEBUFFER -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1908s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	307	100.0	426	3 AAC00681	Human sec
2	307	100.0	470	3 AAK26982	Human bon
3	307	100.0	470	6 ABO01569	Human gen
4	307	100.0	477	4 ABA43588	Human bre
5	307	100.0	477	4 AAK27758	Human bon

6	307	100.0	477	6 ABO02210	Human gen
7	307	100.0	539	5 ABV28618	Human pro
8	307	100.0	539	5 ABV22791	Human pro
9	307	100.0	825	2 AAO66888	Human lym
10	307	100.0	825	2 ACF35988	Human TCR
11	307	100.0	1027	4 AAF56401	Human TCR
12	307	100.0	1027	10 ADB75580	Prostate
13	307	100.0	1027	13 ADP54862	Human PRO
14	307	100.0	1316	5 ABV26485	Human pro
15	307	100.0	1316	5 ABV25093	Human pro
16	307	100.0	1395	12 ADL06443	Human tum
17	307	100.0	1418	6 AAD38828	Human PSN
18	307	100.0	2658	10 ADE08798	Novel DNA
19	304	99.4	486	5 ABV25013	Human pro
20	302	98.4	316	4 ABA47985	Human bre
21	302	98.4	316	4 AAK40015	Human bon
22	302	98.4	316	6 ABO14101	Human gen
23	298	97.1	569	5 ABV43475	Human pro
24	294	95.8	1421	6 ABV65007	Lung can
25	294	95.8	1421	13 ADR65880	Human pro
26	294	95.8	1421	13 ADR66783	Human pro
27	294	95.8	1560	10 ADE06987	Novel cod
28	294	95.8	1799	6 ABE23581	Human gen
29	290	94.5	1080	2 AAG37617	TCR gamma
30	290	94.5	2658	10 ADE08798	Novel DNA
31	284	92.5	1046	12 ADP10449	Reference
32	284	92.5	1586	6 ABE64529	Human ben
33	284	92.5	1586	13 ADP54939	Human PRO
34	279	90.9	1080	1 AAN91698	Sequence
35	264	86.0	786	13 ADP55163	Human PRO
36	251	81.8	460	3 AAC04122	Human pro
37	251	81.8	460	3 AAF15623	Human sec
38	245	79.8	1155	5 AAF29889	Human pro
39	240	78.2	1022	5 ABV24072	Human pro
40	240	78.2	1022	5 ABV24124	Human pro
41	240	78.2	1022	5 ABV25823	Human pro
42	240	78.2	1022	5 ABV24034	Human pro
43	240	78.2	1022	5 ABV23986	Human pro
44	240	78.2	1022	5 ABV23991	Human pro
45	240	78.2	1022	5 ABV23991	Human pro

## ALIGNMENTS

RESULT 1	
ID AAC00681	standard; CDNA; 426 BP.
XX	
AC AAC00681;	
XX	
DT 06-OCT-2000	(first entry)
XX	
DE Human secreted protein 5' EST, SEQ ID NO: 679.	
XX	
KW Human, 5' EST; expressed sequence tag; secreted protein; CDNA isolation;	
XX	
OS Homo sapiens.	
XX	
PN BP1033401-A2.	
XX	
PD 06-SEP-2000.	
XX	
PP 21-FEB-2000; 2000EP-00200610.	
XX	
PR 26-FEB-1999; 99US-0122487P.	
XX	
PA (GEST ) GENSET.	
XX	
PI Dumas Mline Edwards J, Duclert A, Giordano J;	
XX	
DR WPI; 2000-500381/45.	
XX	
DR P-PSDB; AAC00681.	

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNA and genomic DNA that correspond to 5' ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX  
PS Claim 1; SEQ ID NO 679; 71pp + Sequence Listing; English.  
XX  
CC The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. An ORF has been identified within the  
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs  
CC derived from 30 different tissues. EST sequences usually correspond  
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are  
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not  
CC well suited for isolating cDNA sequences derived from the 5' ends of  
CC mRNAs and even in those cases where longer cDNA sequences have been  
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from  
CC mRNAs with intact 5' ends and can therefore be used to obtain full length  
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
CC gene therapy and chromosome mapping procedures. They are used to obtain  
CC upstream regulatory sequences and to design expression and secretion  
CC vectors  
CC  
SQ Sequence 426 BP; 151 A; 83 C; 88 G; 104 T; 0 U; 0 Other;  
XX  
XX  
XX Alignment Scores:  
XX Pred. No.: 5.17e-34 Length: 426  
XX Score: 307.00 Matches: 58  
XX Percent Similarity: 100.00% Conservative: 0  
XX Best Local Similarity: 100.00% Mismatches: 0  
XX Query Match: 100.00% Indels: 0  
XX DB: 3 Gaps: 0  
XX  
US-10-031-158-14 (1-58) x AAC00681 (1-426)  
QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
DB 94 ATGCGAGTGTTCCTCCCAAGCCCACTATTTCTTCTTCATTCCTGAACAAAGCTCC 153  
QY 21 ArgArgLeuGlnIuHstHrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
DB 154 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAGATAC 213  
QY 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
DB 214 ATTGGCAAGAAAGAGAGCAACAGATTCTGGATCCAGAGAGGGAACACCA 267  
XX  
XX  
XX RESULT 2  
XX ID AAK26982 standard; DNA; 470 BP.  
XX AC AAK26982;  
XX  
XX DT 06-NOV-2001 (first entry)  
XX  
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 1539.  
XX  
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;  
XX microarray; cancer; leukaemia; lymphoma; myeloma; ss.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200157276-A2.  
XX  
XX PD 09-AUG-2001.  
XX  
XX PF 30-JAN-2001; 2001WO-US000668.  
XX  
XX PR 04-FEB-2000; 2000US-0180312P.  
XX 26-MAY-2000; 2000US-0207456P.  
XX 30-JUN-2000; 2000US-00608408.  
XX 03-AUG-2000; 2000US-00632366.  
XX 21-SEP-2000; 2000US-0234687P.  
XX 27-SEP-2000; 2000US-0236359P.  
XX

PR 04-OCT-2000; 2000GB-00024263.  
XX  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX DR WPI; 2001-488900/53.  
XX  
XX DE Human genome-derived single exon nucleic acid probes useful for analyzing  
XX gene expression in human bone marrow.  
XX  
XX PS Example 4; SEQ ID NO 1539; 658bp + Sequence Listing; English.  
XX  
XX CC The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
XX bone marrow. They can be used to measure gene expression in bone marrow  
XX samples, which may enable the improved diagnosis and treatment of cancers  
XX such as lymphoma, leukemia and myeloma. The present sequence is one of  
XX the probes of the invention  
XX  
SQ Sequence 470 BP; 148 A; 102 C; 87 G; 133 T; 0 U; 0 Other;  
XX  
XX  
XX Alignment Scores:  
XX Pred. No.: 5.9e-34 Length: 470  
XX Score: 307.00 Matches: 58  
XX Percent Similarity: 100.00% Conservative: 0  
XX Best Local Similarity: 100.00% Mismatches: 0  
XX Query Match: 100.00% Indels: 0  
XX DB: 4 Gaps: 0  
XX  
US-10-031-158-14 (1-58) x AAK26982 (1-470)  
QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
DB 203 ATGCGAGTGTTCCTCCCAAGCCCACTATTTCTTCTTCATTCCTGAACAAAGCTCC 262  
QY 21 ArgArgLeuGlnIuHstHrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
DB 263 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAGATAC 322  
QY 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
DB 323 ATTGGCAAGAAAGAGAGCAACAGATTCTGGATCCAGAGAGGGAACACCA 376  
XX  
XX  
XX RESULT 3  
XX ID ABS01569 standard; DNA; 470 BP.  
XX AC ABS01569;  
XX  
XX DT 19-AUG-2002 (first entry)  
XX  
XX DE Human genome-derived single exon probe from lung SEQ ID No 1560.  
XX  
XX KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;  
XX chronic obstructive pulmonary disease; interstitial lung disease;  
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;  
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;  
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
XX primary ciliary dyskinesia; pulmonary hypertension;  
XX hyaline membrane disease.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200186003-A2.  
XX  
XX PD 15-NOV-2001.  
XX  
XX PF 30-JAN-2001; 2001WO-US000665.  
XX  
XX PR 04-FEB-2000; 2000US-0180312P.  
XX



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 16, 2005, 06:01:41 ; Search time 1790 Seconds  
(without alignments)  
1570.057 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 307

Sequence: 1 MOWPSPSPFPFLQLKQSS.....RYGKGRATFMDPRGTP 58

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2.1/USPTO\_spool\_p/US10031158/runat.14102005.135739.18841/app.query.fastx\_1.199  
-DB=genmb1 -QMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=bls62 -TRANS=human40.cdi -LIST=45  
-DOCLIGN=200 -THR SCORE=pcpt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTMT=pcpt -NORM=ext -HEAPSIZE=500 -MTLEN=0 -MAXLEN=2000000000  
-USER=US10031158 @CGN 1.1 5600 @runat.14102005.135739.18841 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGESUBSET -NEG SCORES=0 -WAIT -DSPLLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*

1: gb\_ba:\*  
2: gb\_ncg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_gy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	307	100.0	410	9	HUMTCGCD1
2	307	100.0	426	6	AX884816 Sequence
3	307	100.0	426	6	BD024426 Sequence
4	307	100.0	470	6	CQ131517 Sequence

5	307	100.0	470	6	CQ290455 Sequence
6	307	100.0	477	6	CQ132293 Sequence
7	307	100.0	477	6	CQ200006 Sequence
8	307	100.0	477	6	CQ291096 Sequence
9	307	100.0	539	6	CQ490917 Sequence
10	307	100.0	539	6	CQ496769 Sequence
11	307	100.0	723	9	AJ583012 Homo sapi
12	307	100.0	726	9	AJ583012 Homo sapi
13	307	100.0	825	6	189883 Sequence 18
14	307	100.0	825	6	X72500 H.sapiens m
15	307	100.0	958	9	HMTCRGAD
16	307	100.0	1027	6	AX074415 Sequence
17	307	100.0	1027	6	AF151103 Homo sapi
18	307	100.0	1054	9	BC062761 Homo sapi
19	307	100.0	1316	6	CQ493215 Sequence
20	307	100.0	1316	6	CQ494607 Sequence
21	307	100.0	1470	9	BC072387 Homo sapi
22	307	100.0	1530	9	HMTCRGA
23	307	100.0	1825	9	BC072396 Homo sapi
24	307	100.0	1825	2	AC130306 Homo sapi
25	307	100.0	140691	9	AF159056 Homo sapi
26	307	100.0	171816	9	AC006033 Homo sapi
27	304	99.0	400	9	HSTCRG1
28	304	99.0	486	6	CQ493135 Sequence
29	302	98.4	316	6	CQ144550 Sequence
30	302	98.4	316	6	CQ204403 Sequence
31	302	98.4	316	6	CQ302987 Sequence
32	298	97.1	569	6	CS011627 Sequence
33	294	95.8	330	9	HMTCCGC
34	294	95.8	635	9	HSTRC54
35	294	95.8	720	9	HSTRC10
36	294	95.8	1013	9	HMTCCGANA
37	294	95.8	1160	9	HSTRC64
38	294	95.8	1402	9	HMTCRGB
39	294	95.8	1421	6	AX32835 Sequence
40	294	95.8	1421	9	HMTCCGXH
41	294	95.8	1455	9	BC039725 Homo sapi
42	292	95.1	757	9	HMTCCGC
43	290	94.5	201	11	BV205008
44	290	94.5	300	9	HMTCCGC
45	290	94.5	330	9	HMTCCGC

#### ALIGNMENTS

RESULT 1	HUMTCGCD1	410 bp	DNA	linear	PRI 14-SEP-1995
LOCUS	Human T-cell receptor gamma-chain (TCRCG1) gene,				
DEFINITION	Human T-cell receptor gamma-chain (TCRCG1) gene,				
CL-region, exon 1.					
ACCESSION	M14996				
VERSION	M14996.1	GI:339076			
KEYWORDS	C-region; T cell receptor gamma-chain; T-cell receptor; germline.				
SEGMENT	1 of 3				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	Leftfanc,M.P., Forster,A. and Rabblites,T.H.				
JOURNAL	Genetic polymorphism and exon changes of the constant regions of				
MEDLINE	the human T-cell rearranging gene gamma				
PROCID	Proc. Natl. Acad. Sci. U.S.A. 83 (24), 9596-9600 (1986)				
COMMENT	Original source text: Homo sapiens (clone: lambda-D19) DNA.				
FEATURES	Entry revisions for [1] kindly provided by M.-P. Leftfanc,				
SOURCE	22-JUN-1988.				
Location/Qualifiers					
1..410					
/organism="Homo sapiens"					
/db_type="genomic DNA"					
/cdl_xref="taxon:9606"					

exon  
/map="7p15-p14"  
/clone="lambda-D19"  
/cell\_line="D-FL"  
/gene="TCRCG1"  
/note="G00-120-408"  
/number=1

ORIGIN 205 bp upstream of BamHI site.

## Alignment Scores:

Pred. No.: 3.7e-30 Length: 410  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x HUMTCGCD1 (1-410)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20

Db 53 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCTTCAATTGCTGAACAAAGCTCC 112

Qy 21 ArgArgLeuGlnIsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40

Db 113 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 172

Qy 41 IleGlyLyLySArgArgAlaThrArgPheTrpAppProArgArgLyThrPro 58

Db 173 ATTGGCAAGAAAAGAGACACACGATTCGGATCCCAAGAGGGGAAACACCA 226

RESULT 2  
AX884816 426 bp DNA linear PAT 18-DEC-2003

LOCUS Sequence 679 from Patent EP1033401.

AX884816

VERSION AX884816.1 GI:40040355

KEYWORDS

ORGANISM

SOURCE

REFERENCE

AUTHORS

TITLE

JOURNAL

Genet (FR)

Location/Qualifiers

1. 426

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

267..425

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAE98930.1"

/db\_xref="GI:40040356"

/translation="MTKNDITMKFSWLTVPKSLDKHRCIVRHHNNKNGVDRIIFP  
PIKTDVITM"

ORIGIN

Alignment Scores:

Pred. No.: 3.86e-30 Length: 426

Score: 307.00 Matches: 58

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x AX884816 (1-426)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20

Db 94 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCTTCAATTGCTGAACAAAGCTCC 153

Qy 21 ArgArgLeuGlnIsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40

Db 154 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 213

Qy 41 IleGlyLyLySArgArgAlaThrArgPheTrpAppProArgArgLyThrPro 58

Db 214 ATTGGCAAGAAAAGAGACACACGATTCGGATCCCAAGAGGGGAAACACCA 267

RESULT 4  
COL31517 470 bp DNA linear PAT 21-JAN-2004

LOCUS Sequence 1539 from Patent WO0157276.

Db 94 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCTTCAATTGCTGAACAAAGCTCC 153

Qy 21 ArgArgLeuGlnIsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40

Db 154 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 213

Qy 41 IleGlyLyLySArgArgAlaThrArgPheTrpAppProArgArgLyThrPro 58

Db 214 ATTGGCAAGAAAAGAGACACACGATTCGGATCCCAAGAGGGGAAACACCA 267

## RESULT 3

BD024426 426 bp DNA linear PAT 27-AUG-2002

LOCUS Sequence tag and encoded human protein.

BD024426

BD024426.1 GI:22565649

VERSION JP 2001269182-A/672

KEYWORDS

ORGANISM

SOURCE

REFERENCE

AUTHORS

TITLE

JOURNAL

Genet

OS

PN

PD

PF

PI

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 16, 2005, 05:43:34 ; Search time 116 Seconds  
(without alignments)  
208.359 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 307  
Sequence: 1 MOWPPSPLEFFLQLKQS.....RYGKGRATRPMDRGRTP 58

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep:\*

2: /cgn2\_6/ptodata/2/pubppaa/PCT\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB.pep:\*

5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/2/pubppaa/PCT\_NEW\_PUBCOMB.pep:\*

7: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep:\*

9: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*

10: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*

11: /cgn2\_6/ptodata/2/pubppaa/US09C\_PUBCOMB.pep:\*

12: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pep:\*

13: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep:\*

14: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep:\*

15: /cgn2\_6/ptodata/2/pubppaa/US10C\_PUBCOMB.pep:\*

16: /cgn2\_6/ptodata/2/pubppaa/US10D\_PUBCOMB.pep:\*

17: /cgn2\_6/ptodata/2/pubppaa/US10E\_PUBCOMB.pep:\*

18: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep:\*

19: /cgn2\_6/ptodata/2/pubppaa/US11\_PUBCOMB.pep:\*

20: /cgn2\_6/ptodata/2/pubppaa/US11\_NEW\_PUB.pep:\*

21: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep:\*

22: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	307	100.0	58	14	US-10-205-823-405 Sequence 405, App
2	307	100.0	58	18	US-10-756-149-5551 Sequence 5551, App
3	307	100.0	58	20	US-11-051-454-405 Sequence 405, App
4	71.5	23.3	100	15	US-10-424-599-206785 Sequence 206785, App
5	67	21.8	81	16	US-10-425-115-189599 Sequence 189599, App
6	64.5	21.0	65	16	US-10-425-115-345446 Sequence 345446, App
7	63	20.5	86	16	US-10-437-963-186030 Sequence 186030, App
8	63	20.5	934	9	US-09-891-216-112 Sequence 12, Appl
9	63	20.5	934	9	US-09-891-216-115 Sequence 15, Appl
10	63	20.5	934	14	US-10-168-402-2 Sequence 2, Appl
11	63	20.5	934	14	US-10-168-402-2 Sequence 39, Appl

12	62	20.2	83	15	US-10-424-599-176153 Sequence 176153, App
13	60.5	19.7	325	15	US-10-425-114-72973 Sequence 72973, App
14	60.5	19.7	572	15	US-10-424-599-280962 Sequence 280962, App
15	60	19.5	160	16	US-10-437-963-109049 Sequence 109049, App
16	60	19.5	900	15	US-10-282-122A-68765 Sequence 68765, App
17	59	19.2	84	16	US-10-425-115-229977 Sequence 229977, App
18	59	19.2	108	16	US-10-437-963-193767 Sequence 193767, App
19	59	19.2	641	20	US-11-097-143-5949 Sequence 5949, App
20	59	19.2	2176	20	US-10-097-143-10950 Sequence 10950, App
21	58.5	19.1	128	16	US-10-425-115-315195 Sequence 315195, App
22	58.5	19.1	440	16	US-10-425-115-307562 Sequence 307562, App
23	58.5	19.1	527	15	US-10-425-114-59801 Sequence 59801, App
24	58.5	19.1	551	15	US-10-425-114-55527 Sequence 55527, App
25	58	18.9	83	15	US-10-424-599-256212 Sequence 256212, App
26	58	18.9	201	16	US-10-767-701-61188 Sequence 61188, App
27	57.5	18.7	156	16	US-10-437-963-110934 Sequence 110934, App
28	57.5	18.7	651	17	US-10-732-923-19734 Sequence 19734, App
29	57.5	18.7	673	15	US-10-369-493-14690 Sequence 14690, App
30	57.5	18.7	703	15	US-10-369-493-15173 Sequence 15173, App
31	57.5	18.7	1003	16	US-10-437-963-190584 Sequence 190584, App
32	57.5	18.7	200	16	US-10-425-115-299370 Sequence 299370, App
33	56	18.2	78	14	US-10-106-638-5379 Sequence 5379, App
34	56	18.2	361	16	US-10-739-930-5604 Sequence 5604, App
35	56	18.2	361	17	US-10-732-923-2960 Sequence 2960, App
36	56	18.2	459	15	US-10-369-493-19269 Sequence 19269, App
37	56	18.2	459	15	US-10-424-599-170289 Sequence 170289, App
38	55.5	18.1	66	16	US-10-425-115-189569 Sequence 189569, App
39	55.5	18.1	79	15	US-10-424-599-270832 Sequence 270832, App
40	55.5	18.1	79	16	US-10-425-115-284110 Sequence 284110, App
41	55.5	18.1	99	15	US-10-424-599-240053 Sequence 240053, App
42	55.5	18.1	124	15	US-10-424-599-228996 Sequence 228996, App
43	55.5	18.1	141	9	US-09-771-161A-99 Sequence 99, Appl
44	55.5	18.1	153	15	US-10-424-599-166111 Sequence 166111, App

#### ALIGNMENTS

RESULT 1

US-10-205-823-405

Sequence 405, Application US/10205823

Publication No. US20030108963A1

GENERAL INFORMATION:

APPLICANT: Schlegel, Robert

APPLICANT: Monahan, John B.

APPLICANT: Endegge, Wilson O.

APPLICANT: Gannavarapu, Manjula

APPLICANT: Gorbacheva, Bella

APPLICANT: Hoersch, Sebastian

APPLICANT: Kametkar, Shubhangi

APPLICANT: Monsey, Angela M.

APPLICANT: Zhao, Xumel

APPLICANT: Anderson, Dustin

TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF PROSTATE CANCER

TITLE OF INVENTION: THERAPY OF PROSTATE CANCER

FILE REFERENCE: MRI-044

CURRENT APPLICATION NUMBER: US/10/205,823

CURRENT FILING DATE: 2002-07-25

PRIOR APPLICATION NUMBER: 60/307,982

PRIOR FILING DATE: 2001-07-25

PRIOR APPLICATION NUMBER: 60/314,356

PRIOR FILING DATE: 2001-08-22

PRIOR APPLICATION NUMBER: 60/325,020

PRIOR FILING DATE: 2001-09-25

PRIOR APPLICATION NUMBER: 60/341,746

PRIOR FILING DATE: 2001-12-12

PRIOR APPLICATION NUMBER: 60/362,158

PRIOR FILING DATE: 2002-03-05

NUMBER OF SEQ ID NOS: 455

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 405  
LENGTH: 58  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-205-823-405

Query Match 100.0%; Score 307; DB 14; Length 58;  
Best Local Similarity 100.0%; Pred. No. 2.7e-33;  
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQMPPSPPLFFFLQILKSSRLRLEHTFVFLNFSIMLRLYIGKKRRATRFMDPRGTP 58  
Db 1 MQMPPSPPLFFFLQILKSSRLRLEHTFVFLNFSIMLRLYIGKKRRATRFMDPRGTP 58

RESULT 2  
US-10-756-149-5551  
Sequence 5551, Application US/10756149  
Publication No. US20050181375A1  
GENERAL INFORMATION:

APPLICANT: Aiz, Natsasha  
APPLICANT: Zlotnick, Albert  
TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND  
FILE REFERENCE: file  
CURRENT APPLICATION NUMBER: US/10/756,149  
CURRENT FILING DATE: 2004-01-12  
NUMBER OF SEQ ID NOS: 5818  
SOFTWARE: Patent in version 3.2  
SEQ ID NO 5551  
LENGTH: 58  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-10-756-149-5551

Query Match 100.0%; Score 307; DB 18; Length 58;  
Best Local Similarity 100.0%; Pred. No. 2.7e-33;  
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQMPPSPPLFFFLQILKSSRLRLEHTFVFLNFSIMLRLYIGKKRRATRFMDPRGTP 58  
Db 1 MQMPPSPPLFFFLQILKSSRLRLEHTFVFLNFSIMLRLYIGKKRRATRFMDPRGTP 58

RESULT 3  
US-11-051-454-405  
Sequence 405, Application US/11051454  
Publication No. US20050191673A1  
GENERAL INFORMATION:

APPLICANT: Schlegel, Robert  
APPLICANT: Monahan, John E.  
APPLICANT: Endege, Wilson O.  
APPLICANT: Gannavarepu, Manjula  
APPLICANT: Gorbacheva, Bella  
APPLICANT: Hoersch, Sebastian  
APPLICANT: Kametkar, Shubhangi  
APPLICANT: Monsey, Angela M.  
APPLICANT: Glatt, Karen  
APPLICANT: Zhao, Xumel  
APPLICANT: Anderson, Dustin  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
TITLE OF INVENTION: THERAPY OF PROSTATE CANCER  
FILE REFERENCE: MRI-044  
CURRENT APPLICATION NUMBER: US/11/051,454  
CURRENT FILING DATE: 2005-02-04  
PRIOR APPLICATION NUMBER: US/10/205,823  
PRIOR FILING DATE: 2002-07-25  
PRIOR APPLICATION NUMBER: 60/307,982  
PRIOR FILING DATE: 2001-07-25  
PRIOR APPLICATION NUMBER: 60/314,356  
PRIOR FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/325,020

PRIOR FILING DATE: 2001-09-25  
PRIOR APPLICATION NUMBER: 60/341,746  
PRIOR FILING DATE: 2001-12-12  
PRIOR APPLICATION NUMBER: 60/362,158  
PRIOR FILING DATE: 2002-03-05  
NUMBER OF SEQ ID NOS: 455  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 405  
LENGTH: 58  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-11-051-454-405

Query Match 100.0%; Score 307; DB 20; Length 58;  
Best Local Similarity 100.0%; Pred. No. 2.7e-33;  
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQMPPSPPLFFFLQILKSSRLRLEHTFVFLNFSIMLRLYIGKKRRATRFMDPRGTP 58  
Db 1 MQMPPSPPLFFFLQILKSSRLRLEHTFVFLNFSIMLRLYIGKKRRATRFMDPRGTP 58

RESULT 4  
US-10-424-599-206785  
Sequence 206785, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 206785  
LENGTH: 100  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_28754C.1.pep  
US-10-424-599-206785

Query Match 23.3%; Score 71.5; DB 15; Length 100;  
Best Local Similarity 42.0%; Pred. No. 0.16;  
Matches 21; Conservative 4; Mismatches 14; Indels 11; Gaps 2;

Qy 1 MQMPPSPPLFFFLQILKSSRLRLEHTFVFLNFSIMLRLYIGKKRRATRFMDPRGTP 39  
Db 9 LOIFSPDPLFLFLSSLSFLDWSFLERSDLLFMKSYFFRRRFSILLR 58

RESULT 5  
US-10-425-115-189599  
Sequence 189599, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 189599  
LENGTH: 81  
TYPE: PRT  
ORGANISM: Zea mays

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - protein search, using sw model

Run on: October 16, 2005, 05:35:23 ; Search time 55 Seconds  
(without alignments)  
78.721 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 307  
Sequence: 1 MQMPPSPFLFPLQLIKOS.....RYGKKRRATRFWDPRGRTP 58

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/ECTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfillseq.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62.5	20.4	77	4	US-09-248-796A-25511 Sequence 25511, A
2	60.5	19.7	72	4	US-09-328-352-7350 Sequence 7350, Ap
3	60	18.5	904	4	US-09-543-681A-4564 Sequence 4564, Ap
4	56	18.2	542	4	US-09-902-540-16320 Sequence 16320, A
5	55.5	18.1	229	4	US-09-489-039A-9391 Sequence 9391, Ap
6	55.5	18.1	316	1	US-08-403-634-4 Sequence 4, Appl
7	55.5	18.1	316	1	US-08-403-634-31 Sequence 31, Appl
8	55.5	18.1	316	3	US-08-913-441B-4 Sequence 4, Appl
9	55.5	18.1	316	3	US-08-913-441B-31 Sequence 31, Appl
10	55.5	18.1	316	4	US-09-571-985C-4 Sequence 4, Appl
11	55.5	18.1	316	4	US-09-571-985C-31 Sequence 31, Appl
12	55	17.9	403	4	US-09-802-213-5 Sequence 5, Appl
13	55	17.9	414	4	US-09-802-213-3 Sequence 3, Appl
14	54	17.6	354	4	US-09-583-110-4854 Sequence 4854, Ap
15	54	17.6	354	4	US-09-107-433-4701 Sequence 4701, Ap
16	53.5	17.4	338	4	US-09-543-681A-7759 Sequence 7759, Ap
17	53.5	17.4	713	1	US-08-190-802A-63 Sequence 63, Appl
18	53.5	17.4	713	3	US-08-477-346-63 Sequence 63, Appl
19	53.5	17.4	713	3	US-08-473-089-63 Sequence 63, Appl
20	53.5	17.4	713	4	US-08-487-072A-63 Sequence 63, Appl
21	53.5	17.4	713	4	US-09-108-857-3 Sequence 3, Appl
22	53	17.3	79	4	US-09-248-796A-26296 Sequence 26296, A
23	53	17.3	163	4	US-09-270-767-40798 Sequence 40798, A
24	53	17.3	163	4	US-09-328-352-7350 Sequence 7350, A
25	53	17.3	273	3	US-08-936-165A-395 Sequence 395, App
26	52.5	17.1	172	4	US-09-543-681A-7315 Sequence 7315, Ap
27	52.5	17.1	407	4	US-09-328-352-7772 Sequence 7772, Ap

28	52.5	17.1	2618	3	US-09-413-814-28 Sequence 28, Appl
29	52	16.9	559	3	US-09-134-001C-1721 Sequence 3721, Ap
30	52	16.9	574	4	US-09-902-540-11805 Sequence 11805, A
31	52	16.9	920	4	US-09-270-767-43994 Sequence 43994, A
32	52	16.9	897	4	US-09-543-681A-4915 Sequence 4915, Ap
33	51.5	16.8	119	4	US-09-270-767-41178 Sequence 41178, A
34	51.5	16.8	119	4	US-09-270-767-56394 Sequence 56394, A
35	51.5	16.8	178	4	US-09-328-352-42377 Sequence 42377, Ap
36	51.5	16.8	354	1	US-07-759-568-2 Sequence 2, Appl
37	51	16.6	114	4	US-09-489-039A-13491 Sequence 13491, A
38	51	16.6	253	4	US-09-270-767-41520 Sequence 41520, A
39	51	16.6	327	4	US-09-248-796A-16639 Sequence 16639, A
40	51	16.6	458	4	US-09-543-681A-5324 Sequence 6324, Ap
41	51	16.6	488	4	US-09-248-796A-15599 Sequence 15599, A
42	51	16.6	534	4	US-09-270-767-43905 Sequence 43905, A
43	51	16.6	815	4	US-09-489-039A-8577 Sequence 8577, Ap
44	51	16.6	873	4	US-09-543-681A-7315 Sequence 7315, Ap
45	50.5	16.4	61	4	US-09-248-796A-22227 Sequence 22227, A

#### ALIGNMENTS

```

RESULT 1
US-09-248-796A-25511
Sequence 25511, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinibock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA.
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
PRIOR FILING DATE: 1998-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 25511
LENGTH: 77
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-25511

Query Match      20.4%; Score 62.5; DB 4; Length 77;
Beet Local Similarity 38.6%; Pred. No. 0.13;
Matches: 17; Conservative 8; Mismatches 18; Indels 1; Gaps 1;

QY      9 LFFPLQLKOSRRLEHT-FVPLRNFSLMLRYIGKKRRATRFW 51
DB      11 LINFPLKKEKGNLNSNFFPLDFILLPLVLEKKRPIRPF 54

RESULT 2
US-09-328-352-7350
Sequence 7350, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 7350
LENGTH: 72
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-7350

Query Match      19.7%; Score 60.5; DB 4; Length 72;

```

Best Local Similarity 27.5%; Pred. No. 0.23;  
Matches 11; Conservative 17; Mismatches 7; Indels 5; Gaps 2;

QY 13 LQKQSSRLHTFVFL--RNFSLMLRYIGKRRARF 50  
DB 12 MQLFRTSTQLOHAYSFMTQNF---IKFLQGRWSPTF 48

## RESULT 3

US-09-543-681A-4564  
Sequence 4564, Application US/09543681A  
Patent No. 6605709  
GENERAL INFORMATION:

APPLICANT: GARY BRETON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
FILE REFERENCE: 2709.1002-001  
CURRENT FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 4564  
LENGTH: 904  
TYPE: PRT  
ORGANISM: Proteus mirabilis

US-09-543-681A-4564

Query Match 19.5%; Score 60; DB 4; Length 904;  
Best Local Similarity 35.3%; Pred. No. 5.1;  
Matches 18; Conservative 7; Mismatches 16; Indels 10; Gaps 2;

QY 15 LKQSSRLHTFVFLNFSMLRYIGKRRATR-----FMDPRGTP 58  
DB 204 LAGQRAIER--FFRKGNLIALKRLARMAQDVDSQYKFRDSGGTP 251

RESULT 4  
US-09-902-540-16320  
Sequence 16320, Application US/09902540  
Patent No. 6833447  
GENERAL INFORMATION:

APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Wiegand, Roger C.  
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-10(15849)B  
CURRENT APPLICATION NUMBER: US/09/902,540  
CURRENT FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/217,883  
PRIOR FILING DATE: 2000-07-10  
NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 16320  
LENGTH: 542  
TYPE: PRT  
ORGANISM: Myxococcus xanthus  
US-09-902-540-16320

Query Match 18.2%; Score 56; DB 4; Length 542;  
Best Local Similarity 30.5%; Pred. No. 11;  
Matches 18; Conservative 10; Mismatches 23; Indels 8; Gaps 2;

QY 4 FPPSPFLPFLQKQSSRL-----EHTFVFLNFSMLRYIGKRRATR-FMDPR 54  
DB 18 FPPSEIFSTLPRRELKGLGADLGFHTVSVRHSGLLRLHLDIDIDITWEGR 76

RESULT 5  
US-09-489-039A-9391  
Sequence 9391, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:

APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
FILE REFERENCE: 2709.2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 9391  
LENGTH: 229  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae

US-09-489-039A-9391

Query Match 18.1%; Score 55.5; DB 4; Length 229;  
Best Local Similarity 26.4%; Pred. No. 4.7;  
Matches 14; Conservative 9; Mismatches 27; Indels 3; Gaps 1;

QY 4 FPPSPFLPFLQKQSSRLHTFVFLNFS---SIMLRYIGKRRARFMDP 53  
DB 51 FPPSPDYVLNLFRRVKAIRFLYLLKVLKSTLILNYLPTNTTTPFSDP 103

RESULT 6  
US-08-403-634-4  
Sequence 4, Application US/08403634  
Patent No. 5674748  
GENERAL INFORMATION:

APPLICANT: Giordano, Antonio  
TITLE OF INVENTION: NOVEL HUMAN CYCLIN-DEPENDENT  
TITLE OF INVENTION: KINASE-LIKE PROTEINS AND METHODS  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Woodcock, Washburn, Kurtz, Mackiewicz &  
ADDRESSER: No. 56747481s  
STREET: One Liberty Place, 46th floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/403,634  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/208,575  
FILING DATE: 08-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1482  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 316 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-403-634-4

Query Match 18.1%; Score 55.5; DB 1; Length 316;  
Best Local Similarity 35.4%; Pred. No. 6.9;  
Matches 17; Conservative 5; Mismatches 23; Indels 3; Gaps 1;

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 16, 2005, 05:24:13 ; Search time 121 Seconds  
(without alignments)  
245.459 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 307  
Sequence: 1 MQMPPSPFLFLLQLKSS.....RYIGKKRRATRPMDPRGTP 58

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result #No	Score	Query Match	Length	ID	Description
1	66.5	21.7	382	Q652R3	O652R3 Oryza sativ
2	66.5	21.7	735	Q704C6	Q704C6 thermoprote
3	66	21.5	482	Q7MW94	Q7MW94 porphyromon
4	64	20.8	524	Q6OHE6	Q6OHE6 brassica hi
5	63	20.5	524	Q6QHD9	Q6QHD9 brassica ca
6	63	20.5	524	Q6QHE0	Q6QHE0 brassica ol
7	63	20.5	524	Q6QHE1	Q6QHE1 brassica mo
8	63	20.5	524	Q6QHE4	Q6QHE4 brassica in
9	63	20.5	934	Q9HCR9	Q9HCR9 homo sapien
10	63	20.5	935	Q8VID6	Q8VID6 rattus norv
11	63	20.5	988	Q96S76	Q96S76 homo sapien
12	62.5	20.4	779	YU20_ARATH	YU20_ARATH
13	62	20.2	524	Q6OHE7	Q6OHE7 brassica in
14	62	20.2	524	Q6OHE7	Q6OHE7 brassica in
15	62	20.2	579	Q76734	Q76734 dictyosteli
16	61	19.9	298	Q7XIN8	Q7XIN8 lactococcus
17	61	19.9	504	Q13102	Q13102 carassius a
18	60	19.5	511	Q9GHA9	Q9GHA9 flagellaria
19	60	19.5	1199	Q9YA79	Q9YA79 homo sapien
20	59	19.2	637	Q9W4R7	Q9W4R7 dirosophila
21	59	19.2	1463	Q86NZ0	Q86NZ0 dirosophila
22	59	19.2	2243	Q9VGP1	Q9VGP1 dirosophila
23	58.5	19.1	1876	Q7RIX5	Q7RIX5 plasmodium
24	58	18.9	253	Q04021	Q04021 arabidopsis
25	58	18.9	374	Q9KCM2	Q9KCM2 bacillus ha
26	58	18.9	622	Q84TU9	Q84TU9 arabidopsis
27	58	18.9	843	Q8SB10	Q8SB10 arabidopsis
28	57.5	18.7	53	Q8FSU4	Q8FSU4 leprospira
29	57.5	18.7	211	YAN8_SCHPO	YAN8_SCHPO
30	57.5	18.7	651	Q7D2Y4	Q7D2Y4 schizosacch
31	57.5	18.7	713	Q8U0I6	Q8U0I6 agrobacteri

32	57.5	18.7	1082	2	Q80U35	Q80U35 mus musculu
33	57	18.6	243	2	Q8A9M4	Q8A9M4 bacteroides
34	57	18.6	492	2	Q8A9R3	Q8A9R3 arabidopsis
35	56.5	18.4	124	2	Q9YR31	Q9YR31 aeropyrum p
36	56.5	18.4	282	2	Q94BB8	Q94BB8 kochia scop
37	56.5	18.4	332	2	Q61982	Q61982 caenorhabdi
38	56.5	18.4	511	1	MARK_BROIN	MARK_BROIN
39	56.5	18.4	821	2	Q73YX5	Q73YX5 bromus iner
40	56	18.2	61	2	Q8V7B3	Q8V7B3 mycobacteri
41	56	18.2	276	2	Q8BYB4	Q8BYB4 tt virus. o
42	56	18.2	344	2	Q7P382	Q7P382 mus musculu
43	56	18.2	361	2	Q9SN11	Q9SN11 fusbacteri
44	56	18.2	563	2	Q8ENR2	Q8ENR2 arabidopsis
45	56	18.2	586	2	Q7WZ14	Q7WZ14 oceanobacil
						Q7WZ14 haemophilus

ALIGNMENTS

RESULT 1	ID	Q652R3	PRELIMINARY;	PRT;	382 AA.
AC	Q652R3				
DT	25-OCT-2004 (TREMBlrel. 28, Created)				
DT	25-OCT-2004 (TREMBlrel. 28, Last sequence update)				
DT	25-OCT-2004 (TREMBlrel. 28, Last annotation update)				
DE	Hypothetical protein P0603C10.53.				
GN	Name=P0603C10.53;				
OS	Oryza sativa (japonica cultivar-group).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;				
OC	Ehrhartoideae; Oryzoideae; Oryza.				
OX	NCBI_TaxID=39947;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Sasaki T., Matsunoto T., Katayose Y.;				
RT	"Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC				
RT	clone: P0603C10."				
RT	Submitted (JUL-2002) to the EMBL/GenBank/DBSJ databases.				
DR	EMBL; AP005527; BAD46204.1; -				
KW	Hypothetical protein.				
SQ	SEQUENCE 382 AA; 41577 MW; 3CAD651AE25B8DBD CRC64;				
Query Match	21.7%; Score 66.5; DB 2; Length 382;				
Best Local Similarity	41.5%; Pred. No. 4.7;				
Matches	17; Conservative 3; Mismatches 16; Indels 5; Gaps 1;				
QY	18 QSSRLHTFYFLRNFSIMLRYIGKKRRATRPMDPRGTP 58				
DB	226 RASRRRHGAVALHNFSRG-----GDRRRRRRHWPFSRRP 261				
RESULT 2	ID	Q704C6	PRELIMINARY;	PRT;	735 AA.
AC	Q704C6				
DT	05-JUL-2004 (TREMBlrel. 27, Created)				
DT	05-JUL-2004 (TREMBlrel. 27, Last sequence update)				
DT	05-JUL-2004 (TREMBlrel. 27, Last annotation update)				
DE	Trehalose-6-phosphate synthase-phosphatase.				
GN	Name=tpsp;				
OS	Thermoproteus tenax.				
OC	Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;				
OC	Thermoproteaceae; Thermoproteus.				
OX	NCBI_TaxID=2271;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	PubMed=15028704; DOI=10.1128/JB.186.7.2179-2194.2004;				
RA	Stiebert B., Tjaden B., Michalke K., Doerr C., Zaparty M.,				
RA	Gordon P., Sengen C.W., Zibat A., Klenk H.P., Schuster S.C.,				
RA	Hensel R.;				
RT	"Reconstruction of the central carbohydrate metabolism of				
RT	Thermoproteus tenax using genomic and biochemical data."				

```

RL J. Bacteriol. 186:2179-2194(2004).
DR EMBL: A521287; CAP18468.1; -.
DR GO: GO:0003825; F:alpha,alpha-trehalose-phosphate synthase (U. .; IEA.
DR GO: GO:0003824; F:catalytic activity; IEA.
DR GO: GO:0004805; F:trehalose-phosphatase activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR GO: GO:0005992; P:trehalose biosynthesis; IEA.
DR InterPro: IPR001830; Glyco_trans_20.
DR InterPro: IPR006379; HAD_SF_11B.
DR InterPro: IPR003357; Trehalose_Phase.
DR Pfam: PF00982; Glyco_transf_20; 1.
DR Pfam: PF02358; Trehalose_Phase; 1.
DR TIGRfam: TIGR01484; HAD_SF_11B; 1.
DR TIGRfam: TIGR00685; T6Pp; 1.
SQ SEQUENCE 735 AA; 82040 MW; AF7A6B83844804ED CRC64;

Query Match
Best Local Similarity 21.7%; Score 66.5; DB 2; Length 735;
Matches 20; Conservative 7; Mismatches 11; Indels 15; Gaps 3;

QY 4 PPPPLFFFLQLLKSSRC--IE-----HTVFRLNFSIMLRYIGKK 44
DB 173 PEPABL---LQLLPSEMRREILGLGSDLVGFHYEYSANFSRSVRLGYK 222

RESULT 3
QY 07MV94 PRELIMINARY; PRT; 482 AA.
AC 1 Q7MV94;
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS OrderedLocustNames=PG1186;
GN Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteri; Bacteroidetes; Bacteroides (Class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=W83;
RX MEDLINE=22829867; PubMed=12949112;
RX DOI=10.1128/JB.185.18.5591-5601.2003;
RA Nelson K.E., Fleisemann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
RA Bisen U.A., Daugherty S.C., Dodson R.U., Durkin A.S., Swinn M.L.,
RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
RA Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
RT Porphyromonas gingivalis strain W83."
RL J. Bacteriol. 185:5591-5601(2003).
DR EMBL: AEO17176; AAC6284.1; -.
DR TIGR: PG1186; -.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 482 AA; 53698 MW; 4E7B69ECDDB6016 CRC64;

Query Match
Best Local Similarity 21.5%; Score 66; DB 2; Length 482;
Matches 12; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 15 LTKSSRLHTEFVFLNFSIMLRY 40
DB 325 LTKSIDREVSHPASLNRYDLITLY 350

RESULT 4
QY 06QHE6 PRELIMINARY; PRT; 524 AA.
AC 06QHE6;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Maturase K.

```

```

GN Name=matK;
OS Brassica hirsutis.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=69182;
RN (1)
RP SEQUENCE FROM N.A.
RA Simonetti E., Martin J.P., Gonzalez L.M., Aguinagalde I.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY541612; AAS48145.1; -.
DR GO: GO:0009507; C:chloroplast; IEA.
DR GO: GO:0008380; P:RNA splicing; IEA.
DR InterPro: IPR008998; Agglutinin.
DR InterPro: IPR004442; Intron_mature2.
DR InterPro: IPR002866; MatK_N.
DR Pfam: PF01348; Intron_mature2; 1.
DR Pfam: PF01824; MatK_N; 1.
KM Chloroplast.
SQ SEQUENCE 524 AA; 62591 MW; 52940B1F109FCEA6 CRC64;

Query Match
Best Local Similarity 20.8%; Score 64; DB 2; Length 524;
Matches 22; Conservative 9; Mismatches 8; Indels 30; Gaps 4;

QY 9 LFFFLQLLKSSR-----RLHTEF-VFLNFSIML-----LRY 40
DB 241 IFFFLR--KQSHLRSTSYDVFFERILFYGKIQHFYKFINNFSLGLKDPFLHYRY 298

```

```

QY 41 IGKRRATR 49
DB 299 HGKYLATK 307

RESULT 5
QY 06QHD9 PRELIMINARY; PRT; 524 AA.
AC 06QHD9;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Maturase K.
GN Name=matK;
OS Brassica campestris (Field mustard).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN (1)
RP SEQUENCE FROM N.A.
RA Simonetti E., Martin J.P., Gonzalez L.M., Aguinagalde I.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY541619; AAS48152.1; -.
DR GO: GO:0009507; C:chloroplast; IEA.
DR GO: GO:0008380; P:RNA splicing; IEA.
DR InterPro: IPR008998; Agglutinin.
DR InterPro: IPR004442; Intron_mature2.
DR InterPro: IPR002866; MatK_N.
DR Pfam: PF01348; Intron_mature2; 1.
DR Pfam: PF01824; MatK_N; 1.
KM Chloroplast.
SQ SEQUENCE 524 AA; 62563 MW; 00040E1F09FCEA2A CRC64;

Query Match
Best Local Similarity 20.5%; Score 63; DB 2; Length 524;
Matches 22; Conservative 9; Mismatches 8; Indels 30; Gaps 4;

QY 9 LFFFLQLLKSSR-----RLHTEF-VFLNFSIML-----LRY 40
DB 241 IFFFLR--KQSHLRSTSYDVFFERILFYGKIQHFYKFINNFSLGLKDPFLHYRY 298

```



A; Cross-references: UNIPROT:004021; GB:AE005172; NID:g2342677; PIDN:AAB70399.1; GSPDB:G



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: October 16, 2005, 05:22:43 ; Search time 122 Seconds  
(without alignments)  
183.870 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 307  
Sequence: 1 MWMPPSPPLFFFLQLKQSS.....RYIGKRRATRFMPDRGTP 58

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	307	100.0	58	4	AAB66399 Human TCR
2	307	100.0	58	7	ADB75581 Prostate
3	307	100.0	58	8	ADP54863 Human PRO
4	91	29.6	16	4	AAB67199 Human TCR
5	63	20.5	934	4	AAG62677 Human typ
6	63	20.5	935	4	AAG62680 Rat type
7	62.5	20.4	688	3	AAG32065 Arabidops
8	62.5	20.4	688	3	AAG40595 Arabidops
9	62.5	20.4	728	3	AAG32064 Arabidops
10	62.5	20.4	728	3	AAG40594 Arabidops
11	62.5	20.4	764	3	AAG40593 Arabidops
12	62.5	20.4	779	3	AAG32063 Arabidops
13	62	20.2	16	4	AAB67200 D discoid
14	61	19.9	116	4	AAO06173 Human pol
15	60.5	19.7	72	6	ADA36663 Actinotoba
16	60.5	19.7	1855	2	AAV21803 B. subcell
17	60.5	19.7	1855	2	AAV83271 Polypepti
18	60	19.5	85	4	AAO07716 Human pol
19	60	19.5	900	6	ABU40841 Protein e
20	60	19.5	904	7	ADFO4279 Bacterial
21	59	19.2	641	4	ABBS9719 Drosophil
22	59	19.2	2176	4	ABBS9719 Drosophil
23	58	18.9	83	4	AAO09340 Human pol
24	57.5	18.7	135	5	ADG34310 Novel hum
25	57.5	18.7	319	8	ABM82552 Human dia

## ALIGNMENTS

26	57.5	18.7	673	8	ADS25657 Bacterial
27	57.5	18.7	673	8	ADS26140 Bacterial
28	57.5	18.7	703	8	ADS22475 Bacterial
29	57	18.6	134	4	AAO03116 Human pol
30	56.5	18.4	119	4	AAO03520 Human pol
31	56	18.2	78	4	AAO03520 Human pol
32	56	18.2	232	3	AAO07077 Arabidops
33	56	18.2	232	3	AAO07077 Arabidops
34	56	18.2	308	3	AAO07076 Arabidops
35	56	18.2	308	3	AAO07076 Arabidops
36	56	18.2	361	3	AAO07075 Arabidops
37	56	18.2	361	3	AAO07075 Arabidops
38	56	18.2	459	8	ADS30236 Bacterial
39	55.5	18.1	229	7	ABO62874 Klebsiell
40	55.5	18.1	313	8	ADR50796 Human c-b
41	55.5	18.1	314	7	AD163090 Human apo
42	55.5	18.1	314	8	AD057044 Human cyc
43	55.5	18.1	316	2	AAW04872 Phosphory
44	55.5	18.1	316	2	AAW04870 Cyclin-de
45	55.5	18.1	323	8	ADF88822 Human kin

RESULT 1  
ID AAB66399 standard; protein; 58 AA.  
AC AAB66399;  
XX  
DT 12-APR-2001 (first entry)  
XX  
DE Human TCRgamma alternate reading frame protein.  
XX  
KW Human; TARP; prostate cancer; breast cancer; immunotherapy;  
KW T cell receptor gamma alternate reading frame protein; TCRgamma.  
XX  
OS Homo sapiens.  
XX  
FN WO200104309-A1.  
XX  
PD 18-JAN-2001.  
XX  
PF 12-JUL-2000; 2000WO-US019039.  
XX  
PR 13-JUL-1999; 99US-0143560P.  
PR 01-OCT-1999; 99US-0157471P.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Paetan I, Essand M, Lee B, Vasmatazis G, Wolfgang C;  
XX  
DR WPI; 2001-081050/09.  
XX  
DR N-PSDB; AAF56401.  
XX  
PT Isolated T-cell receptor gamma alternate reading frame protein useful for  
PT diagnosing and raising an immune response to prostate cancer and breast  
PT cancer.  
XX  
PS Claim 2; Fig 1; 85pp; English.  
XX  
CC The present invention provides the protein and coding sequences of the  
CC human T cell receptor alternate reading frame protein (TRAP). This  
CC protein is expressed in prostate and breast cancer cells at higher levels  
CC than normal and so can be used in the immunotherapy of these cancers, as  
CC well as their detection and prevention  
XX  
SQ Sequence 58 AA;  
Query Match 100.0%; Score 307; DB 4; Length 58;  
Best Local Similarity 100.0%; Pred. No. 2.6e-36;  
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOWPSPPLFFFLQLOKSSRLRHTFVFLANFSMLRLYIGKKRRATFMDPRGTP 58  
DB 1 MOWPSPPLFFFLQLOKSSRLRHTFVFLANFSMLRLYIGKKRRATFMDPRGTP 58

## RESULT 2

ID ADB75581 standard; prolein; 58 AA.

AC ADB75581;

DT 04-DEC-2003 (first entry)

DE Prostate cancer marker prolein.

KW Prostate; cancer; cytostatic; gene therapy; marker.

OS Homo sapiens.

PN WO2003009814-A2.

PD 06-FEB-2003.

PE 25-JUL-2002; 2002WO-US023913.

PF 25-JUL-2001; 2001US-0307982P.

PR 22-AUG-2001; 2001US-0314356P.

PR 25-SEP-2001; 2001US-0325020P.

PR 12-DEC-2001; 2001US-0341746P.

PR 05-MAR-2002; 2002US-0362158P.

PA (MILL-) MILLENNIUM PHARM INC.

XX Schlegel R, Monahan JF, Endege WO, Gannavarapu M, Gorbacheva B,

XX Hoerh S, Kamatkar S, Monsey AM, Glatt K, Zhao X, Anderson D;

XX WPI; 2003-248033/24.

PT New nucleic acid molecule, useful for diagnosing or treating prostate

XX cancer.

PS Disclosure; SEQ ID NO 405; 99pp; English.

CC The invention relates to newly discovered cancer markers associated with  
CC the cancerous state of prostate cells. Also disclosed is a method of  
CC assessing whether a patient is afflicted with prostate cancer. The method  
CC of the invention involves assessing whether a patient is afflicted with  
CC prostate cancer by comparing the level of expression of a marker in a  
CC patient sample and the normal level of expression of the marker in a  
CC control non-prostate cancer sample, where a significant increase in the  
CC level of expression of the marker in the patient sample and the normal  
CC level indicates that the patient is afflicted with prostate cancer.  
CC Nucleic acids of the invention are useful for diagnosing or treating  
CC prostate cancer, and may be useful in gene therapy. Sequences given in  
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 58 AA:

Query Match 100.0%; Score 307; DB 7; Length 58;

Best Local Similarity 100.0%; Pred. No. 2.6e-36;

Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOWPSPPLFFFLQLOKSSRLRHTFVFLANFSMLRLYIGKKRRATFMDPRGTP 58  
DB 1 MOWPSPPLFFFLQLOKSSRLRHTFVFLANFSMLRLYIGKKRRATFMDPRGTP 58

RESULT 3  
ADP54863

ID ADP54863 standard; prolein; 58 AA.

AC ADP54863;

DT 18-NOV-2004 (first entry)

DE Human PRO protein sequence SEQ ID NO:839.

KW human; PRO; immune related disease; inflammatory immune response;  
KW immune response stimulation; anti-allergic; antianemic; antiarthritic;  
KW antidiabetic; antidiabetic; anti-inflammatory; antiproliferative;  
KW antineoplastic; antineoplastic; CNS; dermatological; gastrointestinal;  
KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;  
KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;  
KW virucide; gene therapy.

OS Homo sapiens.

PN WO2004039956-A2.

PD 13-MAY-2004.

PE 28-OCT-2003; 2003WO-US034381.

PR 29-OCT-2002; 2002US-0422472P.

PA (GENTH) GENTECH INC.

XX Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;

XX Wood WT, Wu TD;

XX WPI; 2004-376182/35.

XX N-PEDB; ADP54862.

PT New PRO polynucleotides and polypeptides, useful in diagnosing

XX and treating an immune related disease, e.g. systemic lupus

XX erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in

XX stimulating an immune response.

PS Claim 1; SEQ ID NO 839; 3009pp; English.

CC The present invention describes an isolated PRO nucleic acid (1). Also  
CC described: (1) a vector comprising (1); (2) a host cell comprising the  
CC vector of (1); (3) a process for producing a PRO polypeptide; (4) an  
CC isolated PRO polypeptide; (5) a chimeric molecule comprising the  
CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an  
CC antibody which specifically binds to a polypeptide of (4); (7) a  
CC composition of matter comprising a polypeptide of (4), an agonist or  
CC antagonist of the polypeptide or an antibody that binds to the  
CC polypeptide in combination with a carrier; (8) an article of manufacture  
CC comprising a container, a label on the container and a composition of  
CC matter of (7); (9) a method of treating an immune related disease in a  
CC mammal; (10) a method for determining the presence of a PRO polypeptide  
CC in a sample suspected of having the polypeptide; (11) a method of  
CC diagnosing an immune related disease or an inflammatory immune response  
CC in mammal; (12) a method of identifying a compound that inhibits or  
CC mimics the activity of or expression of a gene encoding a PRO polypeptide  
CC; and (13) a method of stimulating the immune response in a mammal. The  
CC PRO sequences have anti-allergic, anti-inflammatory, antiproliferative,  
CC antineoplastic, antidiabetic, antineoplastic, antineoplastic,  
CC antineoplastic, antineoplastic, antineoplastic, antineoplastic,  
CC haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,  
CC nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and  
CC virucide activities, and can be used in gene therapy. The nucleic acid  
CC (1) and the encoded polypeptides, compositions, kits and methods are  
CC useful in diagnosing and treating an immune related disease and in  
CC stimulating an immune response. The present sequence represents a human  
CC PRO protein from the present invention.

XX Sequence 58 AA:

Query Match 100.0%; Score 307; DB 8; Length 58;

Best Local Similarity 100.0%; Pred. No. 2.6e-36;